

Seq primer: TF
Class: Shotgun.
Location/Qualifiers

1. .693
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="LANDSBERG ERECTA"
/db_xref="taxon:3702"
/clone="LBERM21"
/clone_lib="LBERG"
/note="Organ: Leaf; Vector: pUC19K; Total genomic DNA was sheared to 0.6-0.8 Kbp before ligation."

ORIGIN

Query Match 18.8%; Score 614.4; DB 28; Length 693;
Best Local Similarity 99.8%; Pred. No. 1.3e-99;
Matches 615; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 829 AGTCAATTTCCAGAGAAACAACAAACAGGAAACATTTGAAAACTTGGGTTT 888
DB 688 AGTCAATTTCCAGAGAAACAACAAACAGGAAACATTTGAAAACTTGGGTTT 629
QY 889 CTCGAATTCGAATCTCTTGAAGCTTGGGGTGTCTACGACCAAGGAGGTACCAAG 948
DB 628 CTCGAATTCGAATCTCTTGAAGCTTGGGGTGTCTACGACCAAGGAGGTACCAAG 569
QY 949 TATCTGGGAGAGAAACAATGTATCCGGAGAGGTAAAGCAATGTCTGGGGCATTTCAAG 1008
DB 568 TATCTGGGAGAGAAACAATGTATCCGGAGAGGTAAAGCAATGTCTGGGGCATTTCAAG 509
QY 1009 CTAACATATCTGTCTGGGAGAGAGCTTGAACAGAAATGATTAACAATTTTGGCAC 1068
DB 508 CTAACATATCTGTCTGGGAGAGAGCTTGAACAGAAATGATTAACAATTTTGGCAC 449
QY 1069 CCCACCTGTATCTGCGCTCTCTTGAAGAGATGGAATTTGGACGAGAGAGTT 1128
DB 448 CCCACCTGTATCTGCGCTCTCTTGAAGAGATGGAATTTGGACGAGAGAGTT 389
QY 1129 CTGCTCAGACACAGCTGTGAGAGAGTTCCGACGAGGAGATGATGATATAGTTT 1188
DB 388 CTGCTCAGACACAGCTGTGAGAGAGTTCCGACGAGGAGATGATGATATAGTTT 329
QY 1189 CTGAGAGAGAGATGATTCGATGCTTTGATGATGATGATGATGATGATGATG 1248
DB 328 CTGAGAGAGAGATGATTCGATGCTTTGATGATGATGATGATGATGATGATG 269
QY 1249 ATTATGACTCGAGTGTGATCAAAAGACCATGATCAAGAAACAAATAGTGTTC 1308
DB 268 ATTATGACTCGAGTGTGATCAAAAGACCATGATCAAGAAACAAATAGTGTTC 209
QY 1309 AAAAGTTCTTGGAGCTTGGATAGCTTGCATCGACGATTAATGAACACAGAGGC 1368
DB 208 AAAAGTTCTTGGAGCTTGGATAGCTTGCATCGACGATTAATGAACACAGAGGC 149
QY 1369 AGTGGCATTTGTCAGCTTGTGAGAAACGAGCTGTGCTCATGATGATGATGATG 1428
DB 148 AGTGGCATTTGTCAGCTTGTGAGAAACGAGCTGTGCTCATGATGATGATGATG 89
QY 1429 CTCTACTAGCTCATGC 1444
DB 88 CTCTACTAGCTCATGC 73

RESULT 2
A0964580/c 613 bp DNA linear GSS 28-JAN-2000
LOCUS LERGX20TF LERGX Arabidopsis thaliana genomic clone LERGX20, genomic
DEFINITION
survey sequence.
ACCESSION A0964580
VERSION A0964580.1 GI:6792281
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana

REFERENCE

1 (bases 1 to 613)
Bell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utecht, T.,
Feldblyum, T., Liang, F., Creasy, T., and Frazer, C.M.
Genomic survey sequencing of Landsberg erecta ecotype of
Arabidopsis thaliana and identification of sequence-based
polymorphisms
Unpublished (2000)

JOURNAL

Contact: Xiaoying Lin
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: ac@tigr.org

COMMENT

For additional information, see <http://www.tigr.org/db/at/ac.html>
Seq primer: TF
Class: Shotgun.

FEATURES

source

1. .613
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="LANDSBERG ERECTA"
/db_xref="taxon:3702"
/clone="LBERG"
/clone_lib="LBERG"
/note="Organ: Leaf; Vector: pUC19K; Total genomic DNA was sheared to 0.4-0.7 Kbp before ligation."

ORIGIN

Query Match 18.6%; Score 608.2; DB 28; Length 613;
Best Local Similarity 99.5%; Pred. No 1.7e-98;
Matches 610; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 608 GTTGGCTTGTATTAATCTCAAGTCAATGTGATTTGATTTGATGATGATGATG 667
DB 613 GTTGGCATGTTAATTAATCTCAAGTCAATGTGATTTGATTTGATGATGATGATG 554
QY 668 GTTCTTGTGCTATAGTTGTTAAATAGATTTCTAGGCTGCTGCTCAATGTTAAGA 727
DB 553 GTTCTTGTGCTATAGTTGTTAAATAGATTTCTAGGCTGCTGCTCAATGTTAAGA 494
QY 728 AAAGACGTCAGAGGTGATATAGGCTGAGTTGAACATGTTGATTCAGATTTGAGAG 787
DB 493 AAAGACGTCAGAGGTGATATAGGCTGAGTTGAACATGTTGATTCAGATTTGAGAG 434
QY 788 GACGAGCTGCTTTCTTCAAGATGATGAGAGAGTGGAGTCAATTTCCAAAGAA 847
DB 433 GACGAGCTGCTTTCTTCAAGATGATGAGAGAGTGGAGTCAATTTCCAAAGAA 374
QY 848 CAAGAAACAAACAGAGAAACCTTCTGAAAAAATTGGGTTCTCAAGATTCGAATCTCTC 907
DB 373 CAAGAAACAAACAGAGAAACCTTCTGAAAAAATTGGGTTCTCAAGATTCGAATCTCTC 314
QY 908 TAGAGCTTGGGGTGTGTCAGAGCAAGGAGAGGTAGCAAGTATCTGGAGAGAAACA 967
DB 313 TAGAGCTTGGGGTGTGTCAGAGCAAGGAGAGGTAGCAAGTATCTGGAGAGAAACA 254
QY 968 TGTATCGGAGAGATGACGCAATGCTGCGGCAATTCAGCTAATATCTGTCGGGG 1027
DB 253 TGTATCGGAGAGATGACGCAATGCTGCGGCAATTCAGCTAATATCTGTCGGGG 194
QY 1028 ACGAGCTTGAAGCAAGATGATTAACAATTTTGGACCCCCCACTGATCTGCCCC 1087
DB 193 ACGAGCTTGAAGCAAGATGATTAACAATTTTGGACCCCCCACTGATCTGCCCC 134
QY 1088 TCTTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1147
DB 133 TCTTTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 74
QY 1148 GCAAGAGTTTCTGACGTGAGATGATGATGATGATGATGATGATGATGATGATG 1207

Db 73 GCAGAGATTTCTGACGTCGAGAGATGATGTGATTAATCTTCTGAGAGAAAGAAATGATTC 14
 Oy 1208 CGATGCTTGGAT 1220
 |||||
 Db 13 CGATGCTTGGAT 1

RESULT 3
 AU236368
 LOCUS
 DEFINITION AU236368 RAF14 Arabidopsis thaliana cDNA clone RAF14-93-K05 5', mRNA sequence.
 ACCESSION AU236368
 VERSION AU236368.1 GI:19875537
 KEYWORDS EST.
 SOURCE Arabidopsis thaliana (chale crees)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 647)
 Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y., and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 Unpublished (2002)
 TITLE Contact: Motoaki Seki
 JOURNAL Plant Functional Genomics Research Group
 COMMENT RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: meeki@rtc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially as reported previously (Seki et al., 1998). cDNA cleaved with BamHI and XhoI was ligated to modified lambdaPhiC-1 vector (Carninci et al., submitted for publication) digested with BamHI and SalI. This clone is in a modified plasmid vector. Please visit our web site (http://www.gsc.riken.go.jp/e/plant/index_e.html) for further details.

FEATURES
 source
 1..647
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF14-93-K05"
 /issue_type="root"
 /lab_host="DH10B"
 /clone_1lb="RAF14"
 /note="Site_1: BamHI, Site_2: SalI"

ORIGIN
 Query Match 16.5%; Score 542; DB 9; Length 647;
 Best Local Similarity 99.5%; Pred. No. 1,le-86;
 Matches 564; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

Db 687 GTTGTAAAAATGAGTTCTAGGGCTGCTCAATGCTTAAGAAAAAGACGTTCAAGGCTGCT 746
 |||||
 Db 83 GTTGTAAAAATGAGTTCTAGGGCTGCTCAATGCTTAAGAAAAAGACGTTCAAGGCTGCT 142
 |||||

Oy 747 TATAGGCTGAGTTGAACAGTTGTTCAAGTTTGGAGGAGAGAGAGCTGCTTCTTCA 806
 |||||
 Db 143 TATAGGCTGAGTTGAACAGTTGTTCAAGTTTGGAGGAGAGAGAGCTGCTTCTTCA 202
 |||||

Oy 807 CAAGATATGAGAGAGAGTGGAGAGTCTATTCCAAGAAAGCAAGAACCAAGCAAGAAC 866
 |||||
 Db 203 CAAGATATGAGAGAGAGTGGAGAGTCTATTCCAAGAAAGCAAGAACCAAGCAAGAAC 262
 |||||

Oy 867 ACTTCTGAAAAAATCTGGGTTTCTCAAGATTCGAATCTCTCTAGAGCTTGGGGTGGTCA 926
 |||||
 Db 263 ACTTCTGAAAAAATCTGGGTTTCTCAAGATTCGAATCTCTCTAGAGCTTGGGGTGGTCA 322
 |||||

Oy 927 CAGCAAGAGAGAGGTAGCAACGTATCTGGAGAGAGAAACAATGATCCGGAGAGTAAAC 986
 |||||
 Db 323 CAGCAAGAGAGAGGTAGCAACGTATCTGGAGAGAGAAACAATGATCCGGAGAGTAAAC 382
 |||||

Oy 987 GGCATATGTCGGGGCATTCAAGCTAACATATCTGGTGGGGAGCAAGCGTTTGAAGAAAG 1046
 |||||
 Db 383 GGCATATGTCGGGGCATTCAAGCTAACATATCTGGTGGGGAGCAAGCGTTTGAAGAAAG 442
 |||||

Oy 1047 TATGATACAACTTTGGGACACCCCACTGTATCTGGCCCTCTTTGGAAGAGAGATGG 1106
 |||||
 Db 443 TATGATACAACTTTGGGACACCCCACTGTATCTGGCCCTCTTTGGAAGAGAGATGG 502
 |||||

Oy 1107 AATTGCAAGGCAAGAGAGAGTTCTGCTCAGACACACAGCTTGACAGATTTCTTCAAGCTG 1166
 |||||
 Db 503 AATTGCAAGGCAAGAGAGAGTTCTGCTCAGACACACAGCTTGACAGATTTCTTCAAGCTG 562
 |||||

Oy 1167 GAGGATATGTGATTAATGCTTCTGAGAAAGAGATATTCGATGCTTTGGATGATCT 1226
 |||||
 Db 563 GAGGATATGTGATTAATGCTTCTGAGAAAGAGATATTCGATGCTTTGGATGATG-TCT 620
 |||||

Oy 1227 GATGAGACCTTGCAAGTATGATTAT 1253
 |||||
 Db 621 GATGAGACCTTGCAAGTATGATTAT 647
 |||||

RESULT 4
 AQ959657/c
 LOCUS LEREM21TF LERE Arabidopsis thaliana genomic clone LEREM21, genomic survey sequence.
 DEFINITION AQ959657
 ACCESSION AQ959657
 VERSION AQ959657.1 GI:6787358
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (chale crees)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 606)
 Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Uterbach, T., Feldblum, T., Liang, F., Creasy, T., and Frazer, C.M.
 Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms
 Unpublished (2000)
 TITLE Contact: Xiaoying Lin
 JOURNAL The Institute for Genomic Research
 COMMENT 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: at@tigr.org
 For additional information, see <http://www.tigr.org/cdb/ac/at.html>
 Seq primer: TF
 Class: Shotgun.

FEATURES
 source
 1..606
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="LANDSBERG ERRECTA"
 /db_xref="taxon:3702"
 /clone="LEREM21"
 /clone_1lb="LERE"
 /note="Organ: Leaf, Vector: pUC19UK; Total genomic DNA was sheared to 0.6-0.8 kbp before ligation."

ORIGIN
 Query Match 16.3%; Score 534; DB 28; Length 606;
 Best Local Similarity 100.0%; Pred. No. 3e-85;
 Matches 534; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 911 AGCTTGGGTGTGTGACGACGAGGAGAGGTAGCAACGTATCTGGAGAGAGAAACAATGT 970
 |||||
 Db 606 AGCTTGGGTGTGTGACGACGAGGAGAGGTAGCAACGTATCTGGAGAGAGAAACAATGT 547
 |||||

```

RESULT 5
A0959659

LOCUS      A0959659          523 bp      DNA      linear      GSS 28-JAN-2000

DEFINITION LEREM21TR LERE Arabidopsis thaliana genomic clone LEREM21, genomic
survey sequence.

ACCESSION  A0959659

VERSION    A0959659.1

KEYWORDS   GI:6787360

SOURCE     GSS.

ORGANISM   Arabidopsis thaliana (chale cress)

            Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
            1 (bases 1 to 523)

REFERENCE  Buell,C.R., Lin,X., Pai,G., Barnstead,M., Bowman,C., Uterbach,T.,
            Feldblum,T., Liang,F., Creasy,T. and Fraser,C.M.
            Genomic survey sequencing of Landsberg erecta ecotype of
            Arabidopsis thaliana and identification of sequence-based
            polymorphisms

COMMENT    Unpublished (2000)

JOURNAL    Contact: Xiaoying Lin
            The Institute for Genomic Research
            9712 Medical Center Dr., Rockville, MD 20850, USA
            Tel: 301 838 0200
            Fax: 301 838 0208
            Email: atetigr.org
            For additional information, see http://www.tigr.org/cdb/at/at.html
            Seq primer: TR

FEATURES
            Class: shotgun.

            location/Qualifiers
                1..523
                /organism="Arabidopsis thaliana"
                /mol_type="genomic DNA"
                /strain="LANDSBERG ERECTA"
                /db_xref="taxon:3702"
                /clone="LEREM21"
                /clone_id="LERE"
                /note="Organ: Leaf; Vector: pUC19K; Total genomic DNA was
                sheared to 0.6-0.8 kbp before ligation."

```

Query Match	15.6%;	Score 511.8;	DB 28;	Length 523;
Best Local Similarity	98.7%;	Pred. No. 2.9e-81;		
Matches 516; Conservative	0;	Mismatches 7;	Indels 0;	Gaps 0;

QY	838	CCAAAGAACAACAACAACCAGAAACACTTCGAAAAAAGCTGGTTTCTCAGAA	897
Db	1	CCAAAGAAACMAACAACAACATGAAACACTTCGAAAAAAGCTGGTTTCTCAGAA	60
QY	898	CGAATCTCTAGAGCTTGGGGTGGTCAGCAGCAAGGAGAGAGTAGACAGCATCTGGGA	957
Db	61	CGAATCTCTCTAGAGCTTGGGGTGGTCAGCAACAAGGAGAGTAGACAGCATCTGGGA	120
QY	958	GAGAAACAAATGTAATCCGGGAGAGGTAAACGGCAATGTCGGGGCATTTCAACTAAAT	1017
Db	121	GAGAAACAAATGTAATCCGGGAGAGGTAAACGGCAATGTCGGGGCATTTCAACTAAAT	180
QY	1018	CTGGTCGGGGACGAGCCTTGAGCAGAAAGTAGATTAACAATTGTGCACCCCACCTG	1077
Db	181	CTGGTCGGGGACGAGCCTTGAGCAGAAAGTAGATTAACAATTGTGCACCCCACCTG	240
QY	1078	TATTCGCCCTCTCTTTGGAAAGAGATGCAATTTGGCAGCAAGAGAGTTCTGTCAAC	1133
Db	241	TATTCGCCCTCTCTTTGGAAAGAGATGCAATTTGGCAGCAAGAGAGTTCTGTCAAC	300
QY	1138	ACAAGCTGTGCAGAGATTTCTGACGTGAGAGATGATGTTGGATTAATGCTTCTGAGAA	1197
Db	301	ACAAGATGTGCAGAGATTTCTGACGTGAGAGATGATGTTGGATTAATGCTTCTGAGAA	360
QY	1198	AGATGATTTCCGATGCTTTGGATGATTTCTGATGACGACCTTGCAGATGATTAATGACT	1257
Db	361	AGATGATTTCCGATGCTTTGGATGATTTCTGATGACGACCTTGCAGATGATTAATGACT	420
QY	1258	CGATGTAGTCAAAAAGCCATGATCAAGAAAGCAGAAATAAATGCTTCAAAAAGTCT	1317
Db	421	CGATGTAGTCAAAAAGCCATGATCACTAAAGCAGACTAAATGCTTCAAAAAGTCT	480
QY	1318	TTGGCAGCTTGATAGCTTGTCCATTCGACAGATTAATGAACC	1360
Db	481	TTGGCAGCTTGATAGCTTGTCCATTCGACAGATTAATGAACC	523

RESULT 6	CNS00P9H	457 bp	DNA	linear	GSS 28-JUN-1999
LOCUS	Arabidopsis thaliana genome survey sequence T7 end of BAC F8G21 of				
DEFINITION	ICP library from strain Columbia of Arabidopsis thaliana, genomic				
	survey sequence.				
ACCESSION	AL084227				
VERSION	AL084227.1	GI:5285367			
KEYWORDS	GSS.				
SOURCE	Arabidopsis thaliana (thale cress)				
ORGANISM	Arabidopsis thaliana				
	Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;				
	rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
REFERENCE	1 (bases 1 to 457)				
AUTHORS	Salanoubat,M., Choisme,N., Attiguenave,F., Brottier,P., Wincker,P.,				
	Samson,D., Saurin,W., Weissenbach,J. and Quetier,F.				
JOURNAL	Unpublished				
REFERENCE	2 (bases 1 to 457)				
AUTHORS	Genoscope.				
TITLE	Direct Submissiion				
JOURNAL	Submitted (25-JUN-1999) Genoscope - Centre National de Sequencage				
	BP 101 91006 EVRY cedex - FRANCE (E-mail : segrat@genoscope.cns.fr				
	- Web : www.genoscope.cns.fr)				
FEATURES	Location/Qualifiers				
source	1..457				

```
/organism="Arabidopsis thaliana"  
/mol_type="genomic DNA"  
/strain="Columbia"  
/db_xref="taxon:3702"
```



```

/clone="F8G21"
/clone_lib="IGF"
/note="end : T7"
ORIGIN

```

Query Match	13.8%;	Score 451.2;	DB 29;	Length 457;
Best Local Similarity	99.1%;	Pred. No. 2e-70;		
Matches 453;	Conservative 0;	Mismatches 4;	Indels 0;	Gaps 0;

Qy	1676	TTTACCTCTTAAATTTTCTCTTGATCTCTAGTCTTAGAAGTTTCATTTGATGCGC	1735
Db	1	TTACCTCTTAAATTTTCTCTTGATCTCTAGTCTTAGAAGTTTCATTTGATGCGC	60
Qy	1736	TCGGCATBGCACCAAGAGCTGCTGGAATCTTGACAAAGTATGAGCTCTTAGAGCAC	1795
Db	61	TCGGCATBGCACCAAGAGCTGCTGGAATCTTGACAAAGTATGAGCTCTTAGAGCAC	120
Qy	1796	GCCATTCCTATAGTGTCCACAGGGCCATCGTGGGATGAGTGTTCTGATGTTGAGACAGT	1855
Db	121	GCCATTCCTATAGTGTCCACAGGGCCATCGTGGGATGAGTGTTCTGATGTTGAGACAGT	180
Qy	1856	CCACTGCGTATTTGGAGGCGCCGACGCTCCACCGGAGTTGACTGAGATGGGGTTAGATA	1915
Db	181	CCACTGCGTATTTGGAGGCGCCGACGCTCCACCGGAGTTGACTGAGATGGGGTTAGATA	240
Qy	1916	GAATTCCTCGGGGTCCAGAGCGCAATATGTTTCTGAGGTTGCCAATCTATGGCT	1975
Db	241	GAATTCCTCGGGGTCCAGAGCGCAATATGTTTCTGAGGTTGCCAATCTATGGCT	300
Qy	1976	TCCTTGACAGAGCAAGATCTGGACATTTCAATCAACACTCTCAAGTTCCTCCGCC	2035
Db	301	TCCTTGACAGAGCAAGATCTGGACATTTCAATCAACACTCTCAAGTTCCTCCGCC	360
Qy	2036	AAAGAAATTTGATATATGCTTTAGTTTGTTCATTTGAATTTAAAGTTTGTGGTCCGT	2095
Db	361	AAAGAAATTTGATATATGCTTTAGTTTGTTCATTTGAATTTAAAGTTTGTGGTCCGT	420
Qy	2096	GTTAATGACATCTGTAATATATATATCTATGATTCATT	2132
Db	421	GTTAATGACATCTGTAATATATATATCTATGATTCATT	457

RESULT 7	BH535855	697 bp	DNA	linear	GSS 14-DEC-2001
LOCUS	BH535855				
DEFINITION	BH535855 BOHK Brassica oleracea genomic clone BOHKR13, genomic survey sequence.				
ACCESSION	BH535855				
VERSION	BH535855.1	GI:17767651			
KEYWORDS	GSS.				
SOURCE	Brassica oleracea				
ORGANISM	Brassica oleracea				
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Euarysta; Viridiplantae; Magnoliophyta; eudicotyledons; core eudicots; rosoids; eurosids II; Brassicales; Brassicaceae; Brassica.				
REFERENCE	1 (bases 1 to 697)				
AUTHORS	Town,C.D., Van Aken,S., Uterback,T., Koo,H. and Frazer,C.M.				
TITLE	Whole genome shotgun sequencing of Brassica oleracea				
JOURNAL	Unpublished (2001)				
COMMENT	Other_GSSs: BOHKR13TR				

TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtow@tigr.org
DNA is from a doubled haploid provided by Tom Osborn
Seq primer: TP
Class: sheared ends.
location/Qualifiers
1..697
/organism="Brassica oleracea"
/mol_type="genomic DNA"

```

/strain="T010000D3"
/db_xref="taxon:3712"
/clone="BOHCR3"
/clone_1id="BOHK"
/mote="Vector: pHOSt1, Site 1: BstXI, 2-3 kb sheared
genomic DNA inserted into pHOSt1 using BstXI linkers

```

Query Match	13.7%;	Score 449.2;	DB 28;	Length 697;
Best Local Similarity	80.2%;	Pred. No. 3.9e-70;		
Matches 558;	Conservative 0;	Mismatches 123;	Indels 15;	Gaps 2;

Oy	1302	TGTTTCAAAAAGTCTTTGGCAGCTTGGAAGTCTTGCTCGATCGAGCAGATTAAGAACA	1361
Db	1	TGTTTAAAGAGTTTTCGACAGCTTGGATAGCTTGTCTATCGACGAGATTAAGAGCA	60
Oy	1362	CAGAGCAGTGGCATTGTCAGCTTGTCAGACGCACTTGTCATCGATTGGTATAC	1421
Db	61	CAGAGGCAGTGGCAATTTGTCGGCTTGGCAGAACGGGCTGGCGGTATTTGACTGGTACAA	120
Oy	1422	CTGCACCCCTCACTAGCTCATGCGACGACAAAGAGACTAGCGAGTTAAGCTCCATAGA	1481
Db	121	CTTAGGCACCTTGTAGCTCACGCGAGACAAAGGAGACTAGACGTGTTAAGCTCCACAG	180
Oy	1482	GAATTGGCTGAAGTTTAAAGAAAGATCTACAGATGAGAGCGCATCTGTCAATCTTGT	1541
Db	181	GAATGGCTGAGGTTCTGACAGAGGACTGTGACATGTAGAGGGAGCATCTGTCAATCCCTGC	240
Oy	1542	GGTACAGTTTATGGGCAAGTGGAAAGGTTGGGTGAGATGAAGATTAAGTAAATGTC	1601
Db	241	GGTAAATTTACGGGCAAGTGGAAAGGTTGGGTGAGAGGAAAGACATGATATTGTC	300
Oy	1602	TGGCCTCCAAATGGTCATCATCATGAATACTAGACTGATTAAGACGATTAAGATAGGTG	1661
Db	301	TGGCGCCAAATGGTTCATCATCATGAATACTAGAGCTGACAAAGACGAGAAAGACAAGTT	360
Oy	1662	GAATTTCTTGTCTTTTACTTCTTAAT-----TTTCTTTCGATTTCACTGAT	1711
Db	361	AACCTTCTCTTGTCTTTTCTTTTCTTTTAAATCACTTTGCTTGGCTTGTGCTGAT	420
Oy	1712	-----CTTAGATGTTCATTTGTGTGGCTCGCGATGGGCAACCAAGAGCTGTGTAATA	1766
Db	421	CATAACTTAATGTTGTACATATCATGTGGCTAGAGAAATGGGCAACCGGAGCTGTGAGTA	480
Oy	1767	CTTGCACAGATGAGGCTCTTAGAGACCGCATTTCTTAATGTTCAACAGGCGCATCTGGT	1826
Db	481	CTTTAAAGAGATCTCTCTATTAAAGACCGCATCTCTTAATGTTCAACAGGCGCATCTGGT	540
Oy	1827	GATGAGTGTCTGATGTTTGTAGACATGCGCATGCTAATTTGGAGGCGCGACGCTCCA	1886
Db	541	GATTAGTGTCTGATCTTTGAGAGTACCGCGCATCTGGCTACTTTTAAAGCTACCGGCTACA	600
Oy	1887	CCGGGAGTTGAGTGAATGGGGTTAAGTAAAGATTGCTCTGGGCTCAAGAACCGCATATGTT	1946
Db	601	CAGGAGCTTGAATCAAAAAGGGGTTTGAATGAAGACGCTGGGAGCCGTGCGCGCATAGTGT	660
Oy	1947	TTCTGAGAGTTCGCGCACTGATAGGCTTCCTTGC	1982
Db	661	TTCTGAGAGTTCGCGCACTGATAGGCTTCCTTGC	696

RESULT 8	
AI999551/c	
LOCUS	548 bp mRNA linear EST 08-SEP-1999
DEFINITION	701556368 A. thaliana, Columbia Col-0, rosette-3 Arabidopsis thaliana cDNA clone 701556368, mRNA sequence.
ACCESSION	AI999551
VERSION	AI999551.1
KEYWORDS	GI:5846456
SOURCE	EST.
ORGANISM	Arabidopsis thaliana (thale cress)
	Arabidopsis thaliana
	Eudicotae; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)

REFERENCE
AUTHORS
Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D., Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P., Goxe, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D., Tan, R., Rose, M., Warren, B., Ton, B., Kasurty, K., Borillo, C., Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S., Noblga, A., Murry, L., Turner, C., Krikorian, S., Elder, L., and Hanson, D.

TITLE
JOURNAL
Arabidopsis thaliana Gene Expression Microarray
Unpublished (1999)

COMMENT
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES
source
1..548
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/culti_var="Columbia Col-0"
/db_xref="taxon:3702"
/clone="701556368"
/cisse_type="rosette"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, rosette-3"
/note="Vector: pSPORT; Site_1: NotI; Site_2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1:1:1 peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was blunt-ended, ligated to SalI adaptors, digested with NotI, size selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN
Query Match 13.3%; Score 434.2; DB 9; Length 548;
Best Local Similarity 84.0%; Pred. No. 2e-67;
Matches 536; Conservative 0; Mismatches 12; Indels 90; Gaps 1;

D 2488 GCACACGACAGGTTTTCATGATTCATCAACAGATCCATGAAGAAGACGCAAG 2547
548 GCACACGACAGGTTTTCATGATTCATCAACAGATCCATGAAGAAGACGCAAG 489
Q 2548 GAGAGATTTTCGATGTTTCAGCAGACGAGACGTCGCAAGTTTGGCCACGACG 2607
488 GAGAGATTTTCGATGTTTCAGCAGACGAGACGTCGCAAGTTTGGCCACGACG 429
Q 2608 CAGAACATTATCCCTCTAGCAATGACGATGCGGAAAGAGATATGTAATACTAACA 2667
428 CAGAACATTATCCCTCTAGCAATGACGATGCGGAAAGAGATATGTAATACTAACA 390
Q 2668 TAATCCCTCTGCGCTTTTGTGTTTCAAACTAAGATTAATTCGCGTTTGA 2727
389 ----- 390
Q 2728 TTCTTTCGACGAGTGAAGAGTCAAGCTTCATCGAGTTTCAAGAAAGAGATGAG 2787
389 -----AGAGCTGAAGAGATGTCATTCATCGAGTTTCAAGAAAGAGATGAG 339
Q 2788 GAGTTTGTGAAGAGAGAGATGCTGATAAAGATCAAGAGAGAGATGAGACATG 2847
338 GAGTTTGTGAAGAGAGAGATGCTGATAAAGATCAAGAGAGAGATGAGACATG 279
Q 2848 AAGAGAGGATCCGAGAGATTTTGTGATCTGGAAGAAAGATTTGATGAGGCTTTGAA 2907
278 AAGAGAGGATCCGAGAGATTTTGTGATCTGGAAGAAAGATTTGATGAGGCTTTGAA 219
Q 2908 CAGCTCTGTACAAGCATGGGCTTCAATGAAGATGATGAGACAAAACTCTGTACAC 2967

DB 218 CAGCTCATGTACACAGTGGCTTTCACAAATGAAGATGATGACAAAGCTGTGACAC 159
Q 2368 AAGACAGACTAAGTTTCTTTGTTTCTTTGTTGATGTCGGAAGAGATTCGAGA 3027
DB 158 AAGACAGACTAAGTTTCTTTGTTTCTTTGTTGATGTCGGAAGAGATTCGAGA 99
Q 3028 GACTCCATTTAATCTAGACAAATCTAAGAGATTTAATGATTATTCCTCAATTT 3087
DB 98 GACTCCATTTAATCTAGACAAATCTAAGAGATTTAATGATTATTCCTCAATTT 39
Q 3088 TAGTAGAGGATCTAAGAGACATTAAGTCTGTAC 3125
DB 38 TAGTAGAGGATCTAAGAGACATTAATTTCTGTAC 1

RESULT 9
BH618393/c
LOCUS
DEFINITION
SALK_039005 Arabidopsis thaliana TDNA insertion lines Arabidopsis thaliana genomic clone SALK_039005, genomic survey sequence.
ACCESSION
BH618393
VERSION
BH618393.1 GI:18428488
KEYWORDS
GSS.
ORGANISM
Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 443)
Alonso, J.M., Leisse, T.J., Barajas, P., Chen, H., Cheuk, R., Gadrinab, C., Jeske, A., Karnes, M., Kim, C.J., Parker, H., Prednis, L., Shinn, P., Zimmerman, J., and Ecker, J.R.
A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA. This sequence lies within an annotated intron of At5g23570.
Class: TDNA tagged.
Location/Qualifiers
1..443
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone_lib="SALK_039005"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"

ORIGIN
Query Match 12.7%; Score 417.4; DB 28; Length 443;
Best Local Similarity 99.3%; Pred. No. 2.1e-64;
Matches 440; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

D 1694 CTCTTGACATTTACTGATCTTGAATGTTAATGTTAGTGTGCTCGGACGCAACCAAG 1753
DB 443 CTCTTGACATTTACTGATCTTGAATGTTAATGTTAGTGTGCTCGGACGCAACCAAG 384
Q 1754 AGCTGCTGAAGAACTTCGACAGATGAGGCTCTTGAAGCAGCCATTCATGTCAC 1813
DB 383 AGCTGCTGAAGAACTTCGACAGATGAGGCTCTTGAAGCAGCCATTCATGTCAC 324

QY 1814 AGGGCCATCGGAGATGATGTTCTGATGTTTGAGACAGTGCCTGCTATTGGAGG 1873
 DB 323 AGGGCCATCGGAGATGATGTTCTGATGTTTGAGACAGTGCCTGCTATTGGAGG 264
 QY 1874 CCGAAGCCTTCCA-CCG3GAGTGAAGTGGGTTAGATGAATTCCTGGGGTCA 1932
 DB 263 CCGATCCCTCCANCCGGAGTGTAGTGAAGTGGGTTAGATGAATTCCTGGGGTCA 204
 QY 1933 AAGGCGATGATGTTTCTGAGAGTGTGCGCAACTGATGCTTCCCTGCAACGACAA 1992
 DB 203 AAGGCGATGATGTTTCTGAGAGTGTGCGCAACTGATGCTTCCCTGCAACGACAA 144
 QY 1993 GATTCGACATATCAATCAACACTCTCA-AGTTCCTCCGCCAAGAAATTGATATA 2051
 DB 143 GATTCGACATATCAATCAACACTCTCAAGAGTCTCTCCGCCAAGAAATTGATATA 84
 QY 2052 TGCTTTAGTTTGTGATGGAATTTAAAGTTTGTGCTCGTGTATGATGATCTGTTA 2111
 DB 83 TGCTTTAGTTTGTGATGGAATTTAAAGTTTGTGCTCGTGTATGATGATCTGTTA 24
 QY 2112 TGCTTTAGTTTGTGATGGAATTTAAAGTTTGTGCTCGTGTATGATGATCTGTTA 2134
 DB 23 TGCTTTAGTTTGTGATGGAATTTAAAGTTTGTGCTCGTGTATGATGATCTGTTA 1

RESULT 10
 AU227277/c 421 bp mRNA linear EST 23-APR-2002
 LOCUS AU227277 RAF14 Arabidopsis thaliana cDNA clone RAF14-93-K05 3',
 DEFINITION mRNA sequence.
 AU227277
 ACCESSION AU227277.1 GI:19741924
 VERSION
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Large scale analysis of Arabidopsis full-length cDNA
 Unpublished (2002)
 Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: msek@rtr.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified lambda f10-1 vector (Garnini et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pBluescript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further
 details.

FEATURES
 source
 1. 421
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF14-93-K05"
 /issue_type="root"
 /lab_host="DH10B"
 /clone_lib="RAF14"
 /note="Site_1: BamHI; Site_2: SalI"

ORIGIN
 Query Match 11.9%; Score 389; DB 9; Length 421;

Best Local Similarity 100.0%; Pred. No. 2,6e-59;
 Matches 389; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2737 AGAGCTGAGAGATGTCACAGCTTCATCGAGTTTCAGAGAGAGATGAGAGGTTGTG 2796
 DB 390 AGAGCTGAGAGATGTCACAGCTTCATCGAGTTTCAGAGAGAGATGAGAGGTTGTG 331
 QY 2797 GAAGAGAGAGATGTCATGTAAGATCAAGAGAGAGATGAGAGATGAGAGAGAGG 2856
 DB 330 GAAGAGAGAGATGTCATGTAAGATCAAGAGAGAGATGAGAGATGAGAGAGAGG 271
 QY 2857 CATCAGAGAGATGATTTGATCTGAGAGAGAGATTTGATGAGCTTTGAGAGCTCATG 2916
 DB 270 CATCAGAGAGATGATTTGATCTGAGAGAGAGATTTGATGAGCTTTGAGAGCTCATG 211
 QY 2917 TAAAGAGATGCTTCACATGAGATGATGATGAGAGAGAGATGAGAGAGAGAGAGAG 2976
 DB 210 TAAAGAGATGCTTCACATGAGATGATGATGAGAGAGAGATGAGAGAGAGAGAGAG 151
 QY 2977 CTAAAGTTCTTTGTTGCTTTGATGTCGAGAGAGATGAGAGATCTGAGAGCTCCATT 3036
 DB 150 CTAAAGTTCTTTGTTGCTTTGATGTCGAGAGAGATGAGAGATCTGAGAGCTCCATT 91
 QY 3037 TAAATAGAGAGAGATTTAGAGAGATTTATATCTCCATTTTATGATGAGAG 3096
 DB 90 TAAATAGAGAGAGATTTAGAGAGATTTATATCTCCATTTTATGATGAGAG 31
 QY 3097 GATCTAAGAGAGATTTAGAGAGATTTATATCTCCATTTTATGATGAGAG 3125
 DB 30 GATCTAAGAGAGATTTAGAGAGATTTATATCTCCATTTTATGATGAGAG 2

RESULT 11
 CD835509 644 bp mRNA linear EST 10-JUL-2003
 LOCUS CD835509 BN45.045123F011230 BN45 Brassica napus cDNA clone BN45045123, mRNA
 DEFINITION sequence.
 ACCESSION CD835509
 VERSION
 KEYWORDS
 SOURCE EST.
 ORGANISM Brassica napus (rape)
 Brassica napus
 Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
 1 (bases 1 to 644)
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Genoplante.
 Genoplante, a major partnership french program in plant genomics
 Unpublished (2003)
 Contact: Genoplante
 Genoplante
 93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' (http://www.genoplante.com
 and http://genoplante-info.infobiogen.fr).

FEATURES
 source
 1. 644
 Location/Qualifiers
 /organism="Brassica napus"
 /mol_type="mRNA"
 /culivar="Jet neuf"
 /db_xref="taxon:3708"
 /clone="BN45045123"
 /issue_type="seed"
 /clone_lib="BN45"

ORIGIN
 Query Match 11.8%; Score 385.8; DB 14; Length 644;
 Best Local Similarity 74.8%; Pred. No. 8.3e-59;
 Matches 533; Conservative 0; Mismatches 107; Indels 73; Gaps 1;
 QY 1179 GATTAAGCTTCTGAGAGAGATGATTCGATGCTTTGGATGATTCGATGAGAGACTT 1238

[illegible]

Email: asamiznu@kazusa.or.jp, URL: <http://www.kazusa.or.jp/en/plant/>

FEATURES
 source
 1..380
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="APD25d02R"
 /tissue_type="aboveground organs"
 /dev_stage="two to six-week old"
 /clone_1fb="Arabidopsis thaliana aboveground organs two to six-week old"
 /note="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2: XhoI"

ORIGIN

Query Match	11.6%;	Score 378.4;	DB 9;	Length 380;
Best Local Similarity	99.7%;	Pred. No. 2.1e-57;		
Matches 379;	Conservative 0;	Mismatches 1;	Indels 0;	Gaps 0;
QY 971 ATCCGGAGAGATTAACGGCAATGTCGGGGCATTCAAGCTTAACTATCTGTCGGGACG				1030
DB 1 ATCCGGAGAGATTAACGGCAATGTCGGGGCATTCAAGCTTAACTATCTGTCGGGACG				60
QY 1031 AGCGTTAGCAGGAAGATGATATACACTTTGGTGGACCCGCCCTGATCTGGCCCTCC				1090
DB 61 AGCGTTAGCAGGAAGATGATATACACTTTGGTGGACCCGCCCTGATCTGACCTTC				120
QY 1091 TTGTGAAGAGATGATGATTTGGCAGGACAGAGAGAGTTCCTGTCAGACACAGCTGTGCA				1150
DB 121 TTGTGAAGAGATGATGATTTGGCAGGACAGAGAGAGTTCCTGTCAGACACAGCTGTGCA				180
QY 1151 GGAGTTCTCGACCTGAGAGATGATGTGATATGCTTCGTGAGAGAGAAATGATTTCCGA				1210
DB 181 GGAGTTCTCGACCTGAGAGATGATGTGATATGCTTCGTGAGAGAGAAATGATTTCCGA				240
QY 1211 TGCTTTGGATGATCTGATATGACGACCTTGCAAGATGATTAATGATCGATCGATGAGATCA				1270
DB 241 TGCTTTGGATGATCTGATATGACGACCTTGCAAGATGATTAATGATCGATGAGATCA				300
QY 1271 AAAGAGCATGATCAGAAAGCAGAAATTAAGTGTTCMAAAAGTTCCTTTGGCAGCTTGA				1330
DB 301 AAAGAGCATGATCAGAAAGCAGAAATTAAGTGTTCMAAAAGTTCCTTTGGCAGCTTGA				360
QY 1331 TAGCTTGTGATCGACGACGA 1350				
DB 361 TAGCTTGTGATCGACGACGA 380				

RESULT 13
 BX547725 395 bp DNA linear GSS 02-JUL-2003
 LOCUS Arabidopsis thaliana T-DNA flanking sequence GK-547E01-020631,
 DEFINITION genomic survey sequence.
 ACCESSION BX547725
 VERSION BX547725.1 GI:32440534
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 GSS: Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 Strizhov,N., Li,Y., Rosso,M., Viehovever,P., Dekker,K., Saedler,H.
 and Weisshaar,B.
 A pipeline for automated high-throughput generation of PSTs
 (flanking sequence tags) from Arabidopsis thaliana T-DNA
 transformed lines
 unpublished
 2
 Rosso,M., Strizhov,N., Li,Y., Reiss,B., Dekker,K. and Weisshaar,B.
 A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)
 for flanking sequence tag based reverse genetics

JOURNAL Unpublished
REFERENCE 3 (bases 1 to 395)
AUTHORS Strizhov N., Li Y., Rosso, M. and Weishaar, B.
TITLE JOURNAL
COMMENT Submitted (01-JUL-2003) Weishaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50829, Germany
This sequence is recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by the T-DNA. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:
http://www.mpiz-koeln.mpg.de/GABI-Kat/
location/Qualifiers
1..395
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="GK-547E01-020631"
/note="Arabidopsis thaliana T-DNA insertion lines"
/note="PCR was performed on DNA from Arabidopsis thaliana plants (T1) which were transformed with the T-DNA from vector pAC161. The lines contain one or more T-DNA insertions. The DNA fragment(s) resulting from the PCR were directly sequenced to determine the genomic sequence flanking the insertion. Sequences displaying significant similarity to the A. thaliana nuclear genome sequence were processed for submission. T-DNA derived sequences were removed"

ORIGIN
Query Match 11.1%; Score 364.6; DB 29; Length 395;
Best Local Similarity 95.2%; Pred. No. 5.9e-55;
Matches 376; Conservative 0; Mismatches 19; Indels 0; Gaps 0;

QY 114 ACTGAATGAAATTTGGAGTCCAGATCGGAAACAGAGCGCTTTAGAGCTTAATAG 173
DB 1 ACTGAATGAAATTTGGAGTCCAGATCGGAAACAGAGCGCTTTAGAGCTTAATAG 60
QY 174 CTTCCTCATTTGCTCTCTTCTGTCAGTTATTTCTCTCCGAGCTCTGACTCACTA 233
DB 61 CTTCCTCATTTGCTCTCTTCTGTCAGTTATTTCTCTCCGAGCTCTGACTCACTA 120
QY 234 CTCTCATCTCCGGCGCTTTAACTTACGTTCTCCGCTTACTCTGTAGTTTCTGC 293
DB 121 CTCTCATCTCCGGCGCTTTAACTTACGTTCTCCGCTTACTCTGTAGTTTCTGC 180
QY 294 CTTAGAGCTCCGATCGCTCAACCGCATGCTCTGCTCGATTTCTCTTTCTTCGC 353
DB 181 CTTAGAGCTCCGATCGCTCAACCGCATGCTCTGCTCGATTTCTCTTTCTTCGC 240
QY 354 TGGAAAAATGCGCTTAATGTTCTGATTTGCAAGGTTTGTGCTATGGGTTACTTTT 413
DB 241 TGGAAAAATGCGCTTAATGTTCTGATTTGCAAGGTTTGTGCTATGGGTTACTTTT 300
QY 414 CCCATATTTTATAGTTCTTAGTGAAGTACCTGCGCTTACTGTTTGTATTTTG 473
DB 301 CCCATATTTTATAGTTCTTAGTGAAGTACCTGCGCTTACTGTTTGTATTTTG 360
QY 474 TTGTCCTTCAACGCTTATAGTCTGATCGAGTAT 508
DB 361 TTGTCCTTCAACGCTTATAGTCTGATCGAGTAT 395

RESULT 14
AVS66465/c 619 bp mRNA linear EST 07-SEP-2000
LOCUS AVS66465 Arabidopsis thaliana green siliques Columbia Arabidopsis
DEFINITION Thaliana cDNA clone SQ244D06f 3', mRNA sequence.
ACCESSION AVS66465
VERSION AVS66465.1 GI:8737917
KEYWORDS EST.

SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 619)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S., and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation of 12,028 non-redundant expressed sequence tags from normalized and size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20363093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The first Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
location/Qualifiers
1..619
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="SQ244D06F"
/tissue_type="green siliques"
/clone_lib="Arabidopsis thaliana green siliques Columbia"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI, Site_2: XhoI"

ORIGIN
Query Match 9.2%; Score 301; DB 9; Length 619;
Best Local Similarity 72.8%; Pred. No. 1.1e-43;
Matches 513; Conservative 0; Mismatches 0; Indels 192; Gaps 2;

QY 2170 CAAGAGATGTTTAAAGAGAGTGAAGCTGAGCAGATCTCGAGCAATCAGCAGTGAATAC 2229
DB 619 CAAGAGATGTTTAAAGAGAGTGAAGCAGATCTCGAGCAATCAGCAGTGAATAC 560
QY 2230 TTTAAGAACAGCTCTCAAAACAGAACAGACCCCAAGGCTCTTGAAGATCTTGGAA 2289
DB 559 TTTAAGAACAGCTCTCAAAACAGAACAGACCCCAAGGCTCTTGAAGATCTTGGAA 500
QY 2290 ATTATGACGAGAGAGCTGCTGATGAATCGAGAGTAAATGATTCGAGACAGAACT 2349
DB 499 ATTATGACGAGAGAGCTGCTGATGAATCGAGAGTAAATGATTCGAGACAGAACT 440
QY 2350 AAGATGACGATGAACAGAACAGGGAAGATGATTTTCTAGAAAATCAAAACTT 2409
DB 439 AAGATGACGATGAACAGAACAGGGAAG----- 412
QY 2410 GACATTTTGTATTAACCTAGTCAATTTGATTTGATTAATTTGCCAACAACAAACCTGT 2469
DB 411 ----- 412
QY 2470 GGTGTTTGAAGATGATGACAGACAGGTTTTCATGATTCATCAACAGATTCAT 2529
DB 411 -----AGATGATGACAGACAGGTTTTCATGATTCATCAACAGATTCAT 362
QY 2530 GAAAGAGAGACCCCAAGAGAGAAATTTGAGATGTTGACAGACAGAGAACGTGCCAAG 2589
DB 361 GAAAGAGAGACCCCAAGAGAGAAATTTGAGATGTTGACAGACAGAGAACGTGCCAAG 302
QY 2590 GTTGTGGCCAGACAGACAGAACATTAATCCCTCTACGATATGATTCGCCAAGAGAG 2649
DB 301 GTTGTGGCCAGACAGACAGAACATTAATCCCTCTACGATATGATTCGCCAAGAGAG--- 245
QY 2650 TATATGTACTAATCAATATATCCCTGTGCGTTTGTTCCTCAACCTAAGATTAAT 2709
DB 244 ----- 245
QY 2710 GAATTAATCCGTTTGAATTTCTTCGACAGAGCTGAGAAAGTCAAGCTTCATCGAGTTT 2769

Db 244 -----AGAGCTGAGGAGAGTGTCAAGCTTCATCGAGTTT 212

QY 2770 CAAGAGAAAGATGATGAGAGTTTGTGGAGAGAGGGAGATGCTGATTAAGATCAAGAG 2829

Db 211 CAAGAGAAAGATGAGAGAGTTTGTGGAGAGAGGGAGATGCTGATTAAGATCAAGAG 152

QY 2830 AAGAAGATGAGAGATGAGAGAGGATCATCAGAGAGATATTT 2874

Db 151 AAGAAGATGAGAGATGAGAGAGGATCATCAGAGAGATATTT 107

RESULT 15

AQ010650 290 bp DNA linear GSS 29-MAY-1998

LOCUS F27C8TRC IGF Arabidopsis thaliana genomic clone F27C8, genomic

DEFINITION survey sequence.

ACCESSION AQ010650

VERSION AQ010650.1 GI:3165927

KEYWORDS GSS.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

REFERENCE

AUTHORS Rounsley, S.D., Suh, E.J., Wible, C., Golden, K., Shatsman, S., Choi, P., Yu, K., Akinteyo, B., Shen, K., Geonasekaram, S., Miltner, J., Adams, M.D. and Venter, J.C.

A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 4

TITLE A BAC End Sequence Database for Identifying Minimal Overlaps in Arabidopsis Genomic Sequencing. Update 4

JOURNAL Unpublished (1998)

COMMENT Other GSSs: F27C8TRC

CONTACT: Steve Rounsley

Department of Eukaryotic Genomics

The Institute for Genomic Research

9712 Medical Center Dr., Rockville, MD 20850, USA

Tel: 301 838 0200

Fax: 301 838 0208

Email: rounsley@tigr.org

Seq primer: M13 Reverse

Class: BAC ends

High quality sequence stop: 290.

FEATURES

Source

1..290

Location/Qualifiers

/organism="Arabidopsis thaliana"

/mol_type="genomic DNA"

/strain="Columbia"

/db_xref="taxon:3702"

/clone="F27C8"

/sex="hermaphrodite"

/clone_id="IGF"

/note="Vector: BelobACII; Site_1: EcoRI; Site_2: EcoRI; Produced by Thomas Altmann"

ORIGIN

Query Match 8.9%; Score 290; DB 28; Length 290;

Best Local Similarity 100.0%; Pred. No. 1.4e-41;

Matches 290; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1668 TTCTGCTTTTACTTCTTTAATTTTCTCTGCACTTCTACTGATCTTGAAGTTTACATT 1727

Db 1 TTCTGCTTTTACTTCTTTAATTTTCTCTGCACTTCTACTGATCTTGAAGTTTACATT 60

QY 1728 GTAGTGCTGGGCAATGGGCAACGAAGCTGCTGAATTAATTGCAAGTATGAGGCTCT 1787

Db 61 GTAGTGCTGGGCAATGGGCAACGAAGCTGCTGAATTAATTGCAAGTATGAGGCTCT 120

QY 1788 TAGAGCAGGCATTCTATGCTCCACAGGGCCATGCTGAGTGAAGTTTCTGATGTTGA 1847

Db 121 TAGAGCAGGCATTCTATGCTCCACAGGGCCATGCTGAGTGAAGTTTCTGATGTTGA 180

QY 1848 GAGCAGTGCACCTGCTATTGAGGCCGAAAGCTCCACCGGAGTTAGCTGAGATGGG 1907

Db 181 GAGCAGTGCACCTGCTATTGAGGCCGAAAGCTCCACCGGAGTTAGCTGAGATGGG 240

QY 1908 GTTAGATAGAAATTGCCCTGGGGGTGAGAGCGCAGTATGTTTCTGGAGGTG 1957

Db 241 GTTAGATAGAAATGCTGGGGGTGAGAGCGCAGTATGTTTCTGGAGGTG 290

Search completed: April 8, 2004, 16:24:24

Job time: 5740.02 secs

XX	Beclin C,	Elmayer T,	Vaucheret H;
PI	WPI:	2001-159529/16.	
DR	New SGS3 gene from Arabidopsis thaliana,	useful for increasing virus resistance in plants and, when inhibited, for increasing transgene expression.	
XX	Claim 1;	Page 31-32; 36pp;	French.
PS	The present sequence represents the genomic sequence of the Arabidopsis thaliana SGS3 gene. The SGS3 gene is essential for post-transcriptional inactivation (degradation of RNA) and for resistance to viruses.		
XX	Overexpression of SGS3 results in plants with increased resistance to viruses, while inactivation of SGS3 in transgenic plants (e.g. by expressing antisense RNA, by mutation or by homologous recombination) increases the level of the transgene product. This product may e.g. impart resistance (to herbicide, insects or pathogens), alter contents of essential fatty acids or proteins, or is pharmaceutically active, e.g. an immunoglobulin or interferon		
CC	Sequence 3275 BP, 956 A; 561 C; 804 G; 954 T; 0 U; 0 Other;		
SQ	Query Match	100.0%; Score 3275; DB 4; Length 3275;	
	Best Local Similarity	100.0%; Pred. No. 0;	
	Matches 3275; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1	GACAAACAACAATAATTAAGCAATCTGTTTCGTAGCAATAAATAATAGTGGGAACA	60
DB	1	GACAAACAACAATAATTAAGCAATCTGTTTCGTAGCAATAAATAATAGTGGGAACA	60
QY	61	TTAAGTTAGCGAAAAAGAAAAAAAAGTACAAATAATGAATAACAATCAAT	120
DB	61	TTAAGTTAGCGAAAAAGAAAAAAAAGTACAAATAATGAATAACAATCAAT	120
QY	121	GAAAAATTGGAGTCCAGATCGAAAAACGAGGCCGTTTAGAGCTTAATAACTTCTC	180
DB	121	GAAAAATTGGAGTCCAGATCGAAAAACGAGGCCGTTTAGAGCTTAATAACTTCTC	180
QY	181	AATTTGTCCTCTTCGTAGATTATTTCTTCCTCGGAGTCGACACATCTCAC	240
DB	181	AATTTGTCCTCTTCGTAGATTATTTCTTCCTCGGAGTCGACACATCTCAC	240
QY	241	TCCTCGGCGCTTTAAACTTACGTTCTCGCTGTTTACTCTGAATTTTCGCTTAG	300
DB	241	TCCTCGGCGCTTTAAACTTACGTTCTCGCTGTTTACTCTGAATTTTCGCTTAG	300
QY	301	CCTCGCATCGGCTCACCGCATGCATCTTGTCGTGATTTCTTTTCGCGGAAAA	360
DB	301	CCTCGCATCGGCTCACCGCATGCATCTTGTCGTGATTTCTTTTCGCGGAAAA	360
QY	361	ATTGCGCTAATGTTCTCGATTTTCGAGGTTTTTGTCTAATGGGTTACTTTTTCCCTA	420
DB	361	ATTGCGCTAATGTTCTCGATTTTCGAGGTTTTTGTCTAATGGGTTACTTTTTCCCTA	420
QY	421	TTTAAATGTTCTTAGTAACGATACCTCCGCTTACTGTTTTGTCAATTTTGTGCT	480
DB	421	TTTAAATGTTCTTAGTAACGATACCTCCGCTTACTGTTTTGTCAATTTTGTGCT	480
QY	481	TTCAACCGTTTAGTGCATCGAGATATTTACTGTGAANAATCCTTGTTTTTGTTTT	540
DB	481	TTCAACCGTTTAGTGCATCGAGATATTTACTGTGTGAANAATCCTTGTTTTTGTTTT	540
QY	541	TTGTTTCAATAAATCGAATGATCTACCTTTTGCTTTATGTTTGTGAGGCT	600
DB	541	TTGTTTCAATAAATCGAATGATCTACCTTTTGCTTTATGTTTGTGAGGCT	600
QY	601	ATGCGTTGTTGAGCTTTGTAATCTTCAAGTCAATGATGGAATTTGAATTTGGTATG	660
DB	601	ATGCGTTGTTGAGCTTTGTAATCTTCAAGTCAATGATGGAATTTGAATTTGGTATG	660
QY	661	ACTGAGGTTTCTTGTGCTATAGTTGTAAATAATGATTTCTTAGGCTGTCCAATGT	720

Db	661	ACTGTTGGGTTTCTTGTGGCTTAAGTTGTAAATATAGTTCTAAGGCTGGTCCAAAGT	720
QY	721	CTAAGAAAAGACCTTACGGGTGGTTATAGCCCTGAGTTGAACAGTTGGTCAAGTT	780
Db	721	CTAAGAAAAGACCTTACGGGTGGTTATAGCCCTGAGTTGAACAGTTGGTCAAGTT	780
QY	781	TGGCAGGACGAGACTGGCTTCTTCACAAGATGATGAGAGAGTGGAGGTCAATTCCA	840
Db	781	TGGCAGGACGAGACTGGCTTCTTCACAAGATGATGAGAGAGTGGAGGTCAATTCCA	840
QY	841	AGAAACAAGAACCAACCAAGGAAACCTTCTGGAAAACTTGGGTTCTCAGAAATTCGA	900
Db	841	AGAAACAAGAACCAACCAAGGAAACCTTCTGGAAAACTTGGGTTCTCAGAAATTCGA	900
QY	901	ATCTCTTAGAGCTTGGGGTGGTCAGCAGCAAGGAGAGGTAGCAAGCTATCTGGAGAG	960
Db	901	ATCTCTTAGAGCTTGGGGTGGTCAGCAGCAAGGAGAGGTAGCAAGCTATCTGGAGAG	960
QY	961	GAAACAATGTATCCGGGAGAGGTAAACGGCAATNGTCGGGGCAATCAACTAACAATCTG	1020
Db	961	GAAACAATGTATCCGGGAGAGGTAAACGGCAATNGTCGGGGCAATCAACTAACAATCTG	1020
QY	1021	GTCCGGGACGAGCGTTTGACGAAAGTATGATTAACAATTGTGGCAACCCCACTGTAT	1080
Db	1021	GTCCGGGACGAGCGTTTGACGAAAGTATGATTAACAATTGTGGCAACCCCACTGTAT	1080
QY	1081	CTCGCCCTCTTTGGAAAGAGATGGAATTGGCAGGCAAGAGAGGTTCTGCTACAGCA	1140
Db	1081	CTCGCCCTCTTTGGAAAGAGATGGAATTGGCAGGCAAGAGAGGTTCTGCTACAGCA	1140
QY	1141	CAGCTGTACAGAGTTCCTGACGTGAGAGATGATGTGGAATAAGTCTTGAAGAAAGA	1200
Db	1141	CAGCTGTACAGAGTTCCTGACGTGAGAGATGATGTGGAATAAGTCTTGAAGAAAGA	1200
QY	1201	ATGATTCGGATGCTTTGGATGATTTCTGATGACACCTTGCAAGTGATGATTAAGCTCGG	1260
Db	1201	ATGATTCGGATGCTTTGGATGATTTCTGATGACACCTTGCAAGTGATGATTAAGCTCGG	1260
QY	1261	ATGAGATCAAAAGAGCCATGATCAAGAAAGCAATTAAGTGTCAAAAGTCTCTTG	1320
Db	1261	ATGAGATCAAAAGAGCCATGATCAAGAAAGCAATTAAGTGTCAAAAGTCTCTTG	1320
QY	1321	GCAGCTTGATAGCTTGTGCATGACGACGATTAATGAACCAAGAGCAGTGGCAATTGC	1380
Db	1321	GCAGCTTGATAGCTTGTGCATGACGACGATTAATGAACCAAGAGCAGTGGCAATTGC	1380
QY	1381	CAGCTGTCAAAAGAGCTGTGTCCATGATGTATTAACCTGCACCCTCTTACTAGCTC	1440
Db	1381	CAGCTGTCAAAAGAGCTGTGTCCATGATGTATTAACCTGCACCCTCTTACTAGCTC	1440
QY	1441	ATGCGAGACAAAAGAGCTAGGGGAGTTAAGCTCCATAGAGAAATGGCTGAATTTTAA	1500
Db	1441	ATGCGAGACAAAAGAGCTAGGGGAGTTAAGCTCCATAGAGAAATGGCTGAATTTTAA	1500
QY	1501	AAAAGGATCTACAGATGAGAGCGCATCTGTCAATTCCTTGGTGAATTTATGGCAGT	1560
Db	1501	AAAAGGATCTACAGATGAGAGCGCATCTGTCAATTCCTTGGTGAATTTATGGCAGT	1560
QY	1561	GGAAGGTTTGGGTGAGAGATCAAAAGATTAATGAATTTGTGGCTCCCAATGGTCAACA	1620
Db	1561	GGAAGGTTTGGGTGAGAGATCAAAAGATTAATGAATTTGTGGCTCCCAATGGTCAACA	1620
QY	1621	TCATGAATACTAGACTGATTAAGGACGATACGATTAAGGTGAATTTCTTGTCTTTTAC	1680
Db	1621	TCATGAATACTAGACTGATTAAGGACGATACGATTAAGGTGAATTTCTTGTCTTTTAC	1680
QY	1681	TTCTTTAATTTTCTCTGCAATTCCTACGATCTTAAGATGTTAATATGTAGTGGCTGGC	1740
Db	1681	TTCTTTAATTTTCTCTGCAATTCCTACGATCTTAAGATGTTAATATGTAGTGGCTGGC	1740
QY	1741	ATGGGCAACCAAGAGCTGTGGAATACCTTGACAAAGATGAGAGCTCTTAAGACAGCAT	1800

```

Db      1741 ATGGGCAACCAAGAGCTGCTGGAATCTTGACAAGATGAGGCTCTTAGACACGCCAT 1800
Qy      1801 TCCTATGTCACAGAGGSCCATCGGGATGAGTGTCTGATGTTGAGACAGTGCCT 1860
Db      1801 TCCTATGTCACAGAGGSCCATCGGGATGAGTGTCTGATGTTGAGACAGTGCCT 1860
Qy      1861 GGGCTATTGGAGGCCGCAAGCCCTCCACCGGAGTTAGCTGAGATGGGGTTAGTAGAAT 1920
Db      1861 GGGCTATTGGAGGCCGCAAGCCCTCCACCGGAGTTAGCTGAGATGGGGTTAGTAGAAT 1920
Qy      1921 GCGTGGGCTGAGAAAGCGCAGATGTTTCTGAGAGTGTTCGCCACTGATGCTTCTT 1980
Db      1921 GCGTGGGCTGAGAAAGCGCAGATGTTTCTGAGAGTGTTCGCCACTGATGCTTCTT 1980
Qy      1981 GCAACGAGCAAGATCGGACATATTCATCAACCTCAAGGTCCTCCCCCAAGA 2040
Db      1981 GCAACGAGCAAGATCGGACATATTCATCAACCTCAAGGTCCTCCCCCAAGA 2040
Qy      2041 AATTGATATATGCTTTTATGTTTGTGATTTGAAATTTAAAGTTTGTGGTCCGTGTA 2100
Db      2041 AATTGATATATGCTTTTATGTTTGTGATTTGAAATTTAAAGTTTGTGGTCCGTGTA 2100
Qy      2101 TGCATCTGTTATGTAATATCTATGATTCATTAGGCAAAACAAGCTGAATTCAGTTG 2160
Db      2101 TGCATCTGTTATGTAATATCTATGATTCATTAGGCAAAACAAGCTGAATTCAGTTG 2160
Qy      2161 AATTCATACCAAGAGATGTTTAAAGGAGCTGAGGAGATCTCTGAGCAATCCAGCAG 2220
Db      2161 AATTCATACCAAGAGATGTTTAAAGGAGCTGAGGAGATCTCTGAGCAATCCAGCAG 2220
Qy      2221 CTGAATCTATTAAAGAACAGCTCTCAAAACAGAACAGACGCCAAGTGTCTGAGGAA 2280
Db      2221 CTGAATCTATTAAAGAACAGCTCTCAAAACAGAACAGACGCCAAGTGTCTGAGGAA 2280
Qy      2281 TCTCTGGAATTTATGAGCGAGAAAGCTGCGTGAACCTGACAGAGATATCCGATCGTGA 2340
Db      2281 TCTCTGGAATTTATGAGCGAGAAAGCTGCGTGAACCTGACAGAGATATCCGATCGTGA 2340
Qy      2341 CAGAGAATTAAGATGAGCATGAGCAAGCAAGGAAAGAGATGATTTTCTAGAAAT 2400
Db      2341 CAGAGAATTAAGATGAGCATGAGCAAGCAAGGAAAGAGATGATTTTCTAGAAAT 2400
Qy      2401 CACAAATCTGACATTTTGTATTAATCTACTGATTCACATTTTGTATTAATTTGTCAACA 2460
Db      2401 CACAAATCTGACATTTTGTATTAATCTACTGATTCACATTTTGTATTAATTTGTCAACA 2460
Qy      2461 AAAAAGCTGTGTGTTGAAGATGATGACACAGCAAGGTTTTCATGATTCATCAAA 2520
Db      2461 AAAAAGCTGTGTGTTGAAGATGATGACACAGCAAGGTTTTCATGATTCATCAAA 2520
Qy      2521 CAGATCATGAAGAAGAGACGCAAGGAGAAATTTGAGATGTTGACAGCAGAGAA 2580
Db      2521 CAGATCATGAAGAAGAGACGCAAGGAGAAATTTGAGATGTTGACAGCAGAGAA 2580
Qy      2581 CGTGCCAAAGTTGTGGCCAGCAGACAGCAATTAATCCCTAGCAATGACGATTCG 2640
Db      2581 CGTGCCAAAGTTGTGGCCAGCAGACAGCAATTAATCCCTAGCAATGACGATTCG 2640
Qy      2641 CGAAAGAGGTATATGTAATCAATCAATTAATCCCTGCGGTTTGTGTTTCAAACTTA 2700
Db      2641 CGAAAGAGGTATATGTAATCAATCAATTAATCCCTGCGGTTTGTGTTTCAAACTTA 2700
Qy      2701 AGAGTAAGTAATTAATCCGGTTTGAATCTTTCGAGAGCTGAGAGATGATCAAGCTTC 2760
Db      2701 AGAGTAAGTAATTAATCCGGTTTGAATCTTTCGAGAGCTGAGAGATGATCAAGCTTC 2760
Qy      2761 ATCGAGTTTCAAGAGAAAGATGAGAGATTTGTGGAAGAGAGGAGATGCTGATTA 2820
Db      2761 ATCGAGTTTCAAGAGAAAGATGAGAGATTTGTGGAAGAGAGGAGATGCTGATTA 2820
Qy      2821 GATCAAGAGAAAGATGAGAGATGAGAGAGGATCAAGAGAGATTTGATCTG 2880
Db      2821 GATCAAGAGAAAGATGAGAGATGAGAGAGGATCAAGAGAGATTTGATCTG 2880

```

```

Qy      2881 GAGAAAGAAATTTGATGAGGCTTTGGAACAGCTCATGATACAGCAGCCCTCAATGAA 2940
Db      2881 GAGAAAGAAATTTGATGAGGCTTTGGAACAGCTCATGATACAGCAGCCCTCAATGAA 2940
Qy      2941 GATGATTGACAAAAGTCTGGTACACAGACAGACGTAAGTTCTTTGTTTGTCTTTG 3000
Db      2941 GATGATTGACAAAAGTCTGGTACACAGACAGACGTAAGTTCTTTGTTTGTCTTTG 3000
Qy      3001 GTATGTCGAGAAAGTAGAGATCTGAGAGACTCCATTAAATCTAGACAAATCTAAGA 3060
Db      3001 GTATGTCGAGAAAGTAGAGATCTGAGAGACTCCATTAAATCTAGACAAATCTAAGA 3060
Qy      3061 GATTATAGATTTATATCTCCAAATTTTGTAGACGATCTAAGAGATTAAGTCTT 3120
Db      3061 GATTATAGATTTATATCTCCAAATTTTGTAGACGATCTAAGAGATTAAGTCTT 3120
Qy      3121 GTGACTTAAACCAAGTTTCTTAGTAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 3180
Db      3121 GTGACTTAAACCAAGTTTCTTAGTAATTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 3180
Qy      3181 TTGACATATTTACCAAGCTGAGATGATCACAGAAATGCAATCAATCATGTTT 3240
Db      3181 TTGACATATTTACCAAGCTGAGATGATCACAGAAATGCAATCAATCATGTTT 3240
Qy      3241 AGAATTTTATCTACAAATTAATGAGTACAAT 3275
Db      3241 AGAATTTTATCTACAAATTAATGAGTACAAT 3275

RESULT 2
AAF25374
ID   AAF25374 standard; cDNA; 1878 BP.
XX
AC   AAF25374;
XX
DT   15-MAY-2001 (first entry)
XX
DE   Nucleotide sequence of the Arabidopsis SGS3 polypeptide.
XX
KW   SGS3 gene; post-transcriptional inactivation; RNA degradation;
XX   viral resistance; resistance; fatty acid content; protein content; ss.
XX
OS   Arabidopsis thaliana.
XX
FH   Key
XX   CDS      1..1878
XX           /tag= a
XX           /product= "SGS3"
XX
PN   MO200105951-A2.
XX
PD   25-JUN-2001.
XX
PF   13-JUL-2000; 2000MO-FR002052.
XX
PR   16-JUL-1999; 99FR-00009417.
XX
PR   26-JUN-2000; 2000FR-00001006.
XX
XX
PA   (AVET ) AVENTIS CROPS SCIENCE SA.
XX   (INRG ) INST NAT RECH AGRONOMIQUE.
XX
PI   Beclin C, Elmayan T, Vaucheret H;
XX
DR   WPI, 2001-159529/16.
XX
DR   P-PSDB; AAB31798.
XX
XX
PT   New SGS3 gene from Arabidopsis thaliana, useful for increasing virus
XX   resistance in plants and, when inhibited, for increasing transgene
XX   expression.
XX
PS   Claim 1; Page 32-35; 36pp; French.
XX

```

CC The present sequence encodes the Arabidopsis thaliana SGS3 polypeptide.
 CC The SGS3 gene is essential for post-transcriptional inactivation
 CC (degradation of RNA) and for resistance to viruses. Overexpression of
 CC SGS3 results in plants with increased resistance to viruses, while
 CC inactivation of SGS3 in transgenic plants (e.g. by expressing antisense
 CC RNA, by mutation or by homologous recombination) increases the level of
 CC the transgene product. This product may e.g. impart resistance (to
 CC herbicide, insects or pathogens), alter contents of essential fatty acids
 CC or proteins, or is pharmaceutically active, e.g. an immunoglobulin or
 CC interferon

XX Sequence 1878 BP; 575 A; 332 C; 554 G; 417 T; 0 U; 0 Other;

Query Match 44.6%; Score 1462; DB 4; Length 1878;

Best Local Similarity 83.3%; Pred. No. 0;

Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;

QY 696 ATGAGTTCTAGGGCTGCTCCATGTCCTAAGAAAGAAAGCGTTGAGGTGTTAAGGCTT 755
 Db 1 ATGAGTTCTAGGGCTGCTCCATGTCCTAAGAAAGAAAGCGTTGAGGTGTTAAGGCTT 60

QY 756 GAGTTGAACAGTTGGTTCAAGTTGGCAGGAGCAGACTGGCTTCTTCAAGATGAT 815
 Db 61 GAGTTGAACAGTTGGTTCAAGTTGGCAGGAGCAGACTGGCTTCTTCAAGATGAT 120

QY 816 GGAGAGAGTGGAGGTCATTTCCAAAGAAACAAAGCAAGCAAGCAAGCAAGCAAGCA 875
 Db 121 GGAGAGAGTGGAGGTCATTTCCAAAGAAACAAAGCAAGCAAGCAAGCAAGCAAGCA 180

QY 876 AAAAATTGGGTTTCTCAGAAATTCCTCTAGAGCTTGGGTTGTCAGCAGCAAGG 935
 Db 181 AAAAATTGGGTTTCTCAGAAATTCCTCTAGAGCTTGGGTTGTCAGCAGCAAGG 240

QY 936 AAGAGTGAACAGTATCTGGAGAGAAACAAATGATTCGGAGAGAGTAAACCGCATGCT 995
 Db 241 AAGAGTGAACAGTATCTGGAGAGAAACAAATGATTCGGAGAGAGTAAACCGCATGCT 300

QY 996 CGGGGATTCAGAGCTAATATCTGGTCGGGAGCAGAGCTTGAAGCAAGTATGATTAAC 1055
 Db 301 CGGGGATTCAGAGCTAATATCTGGTCGGGAGCAGAGCTTGAAGCAAGTATGATTAAC 360

QY 1056 AACTTTGTGGACCCCACTGTATTCGCTCTCTTGGAGAGAGATGAAATTGGCAG 1115
 Db 361 AACTTTGTGGACCCCACTGTATTCGCTCTCTTGGAGAGAGATGAAATTGGCAG 420

QY 1116 GCAAGAGAGGTCTCTGTCAGCAGAGCTGTGCAAGAGGTTCTGACGTCAGATGAT 1175
 Db 421 GCAAGAGAGGTCTCTGTCAGCAGAGCTGTGCAAGAGGTTCTGACGTCAGATGAT 480

QY 1176 GTGATATATGCTTCTGAGAGAGATGATTCGATGCTTGGATGATTCGATGACAG 1235
 Db 481 GTGATATATGCTTCTGAGAGAGATGATTCGATGCTTGGATGATTCGATGACAG 540

QY 1236 CTTCGAGATGATATATGATCTCGATGTGATCAAAAGAGCCATGATCAAGAAAGCAG 1295
 Db 541 CTTCGAGATGATATATGATCTCGATGTGATCAAAAGAGCCATGATCAAGAAAGCAG 600

QY 1296 AATAAGTGTCAAAAGTCTTGGCAGCTTGGATGATGCTTGGATGATGATGATGATGAT 1355
 Db 601 AATAAGTGTCAAAAGTCTTGGCAGCTTGGATGATGCTTGGATGATGATGATGATGAT 660

QY 1356 GAACCAAGAGCAGTGTGCTGAGCTTGTCAAGACGAGCTGTGATGATGATGATGATGAT 1415
 Db 661 GAACCAAGAGCAGTGTGCTGAGCTTGTCAAGACGAGCTGTGATGATGATGATGATGAT 720

QY 1416 TATTAACCTCAGCTCTTACTAGCTCATGCGAGCAAAAGAGCTGAGTAAAGCTC 1475
 Db 721 TATTAACCTCAGCTCTTACTAGCTCATGCGAGCAAAAGAGCTGAGTAAAGCTC 780

QY 1476 CATAGAGATTTGGCTGAAGTTTGAAGAAAGATCTACATGAGAGCCGATCTGTGAT 1535
 Db 781 CATAGAGATTTGGCTGAAGTTTGAAGAAAGATCTACATGAGAGCCGATCTGTGAT 840

QY 1536 CCTTGGTGAAGATTTATGCGAGTGAAGGGTTTGGTGAAGATGAAAGATATTAGA 1595
 Db 841 CCTTGGTGAAGATTTATGCGAGTGAAGGGTTTGGTGAAGATGAAAGATATTAGA 900

QY 1596 ATTGTGCTCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
 Db 901 ATTGTGCTCCCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960

QY 1656 AAGGTGAGATTTCTGCTTTTACTCTTAAATTTTCTTGAATCTACTGATCTTA 1715
 Db 961 A----- 961

QY 1716 GAATGTTACATTTAGTGGCTGCGATGGCAGCAAGAGCTGTGATATCTTGACA 1775
 Db 962 -----AGTGGCTGCGATGGCAGCAAGAGCTGTGATATCTTGACA 1007

QY 1776 GTATAGGCTCTTAAAGCAGCGCATTTCTATGTCACAGGCGCATCGTGGATGATGT 1835
 Db 1008 GTATAGGCTCTTAAAGCAGCGCATTTCTATGTCACAGGCGCATCGTGGATGATGT 1067

QY 1836 TCTGATGTTTGAAGACAGTGCACCTGGCTATTTGAGAGCCCAAGCCCTCCACCGGAGTT 1895
 Db 1068 TCTGATGTTTGAAGACAGTGCACCTGGCTATTTGAGAGCCCAAGCCCTCCACCGGAGTT 1127

QY 1896 AACTGAGATGGGCTTAAAGATGAAATGCTGGGCTCAGAGCGCATATGTTTCTGAGG 1955
 Db 1128 AACTGAGATGGGCTTAAAGATGAAATGCTGGGCTCAGAGCGCATATGTTTCTGAGG 1187

QY 1956 TGTTCGCAACTGATGAGCTTCTTGAAGCAAGAGATCTGACATATTCATCAACA 2015
 Db 1188 TGTTCGCAACTGATGAGCTTCTTGAAGCAAGAGATCTGACATATTCATCAACA 1247

QY 2016 CTCTGAGGTTCTCTCCCAAGAAATTTGATATATGCTTTTGAATTTGATGATGAT 2075
 Db 1248 CTCTCA----- 1253

QY 2076 TTAAGTTTGTGGTCGATGATGATGATGATGATGATGATGATGATGATGATGAT 2135
 Db 1254 -----AGG 1256

QY 2136 CAAAACAAAGCTGAATTCGATGAAATCATACCAAGATGTTGTAAGAGACTGAG 2195
 Db 1257 CAAAACAAAGCTGAATTCGATGAAATCATACCAAGATGTTGTAAGAGACTGAG 1316

QY 2196 GCGATGCTTGAAGCAGTCAAGAGCTGAACATCTTGAAGCAAGCTCTCAAAACAGAA 2255
 Db 1317 GCGATGCTTGAAGCAGTCAAGAGCTGAACATCTTGAAGCAAGCTCTCAAAACAGAA 1376

QY 2256 CAAAGCAGCAGGCTTGAAGATCTCTGAAATTTGAAGCAAGAGCTGCTGAGAAC 2315
 Db 1377 CAAAGCAGCAGGCTTGAAGATCTCTGAAATTTGAAGCAAGAGCTGCTGAGAAC 1436

QY 2316 TGCAGAGATTAATCGATGCTGAGCAGAGAACTAAGATGCAAGATGAAACAGAGAG 2375
 Db 1437 TGCAGAGATTAATCGATGCTGAGCAGAGAACTAAGATGCAAGATGAAACAGAGAG 1496

QY 2376 AGAGTATGATTTTCTGAAGAAATCAAACTTGACATTTGATTACTACTGATTTCA 2435
 Db 1497 AG----- 1498

QY 2436 CATTTTGTATTAATTTGTCACAAACAAACCTGTGTTGTAAGATGATGACAGCA 2495
 Db 1499 -----AGATGATGTCACAGCA 1514

QY 2496 CAGTTTTTTCATGATTCATCAATCAACAGATTCATGAAAGAGAGAGAGAGAGAGAGAG 2555
 Db 1515 CAGTTTTTTCATGATTCATCAATCAACAGATTCATGAAAGAGAGAGAGAGAGAGAGAG 1574

QY 2556 TTTGAGATGTTTGAAGCAGAGCAAGCAAGTGTGAGCAAGCAGAGCAGAGAGAGAT 2615
 Db 1575 TTTGAGATGTTTGAAGCAGAGCAAGCAAGTGTGAGCAAGCAGAGCAGAGAGAGAT 1634

QY 2616 TAATCCCTTAGCATGACGATGCGCAAGAGGATATGATTAATCAATATCCCT 2675

```

Db      1635 TATCCCTAGCAATGACGATTGCCGAAG----- 1665
QY      2676 CTGGCGTTTGTGTTTCAAACTAAGATTAATTCGCTTTGATTCCTTCG 1735
Db      1666 ----- 1665
QY      2736 CAGAGCTGAGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGATTTGT 2795
Db      1666 -AGAGCTGAGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGATTTGT 1724
QY      2796 GGAAGAGAGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGATTTGT 2855
Db      1725 GGAAGAGAGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGATTTGT 1784
QY      2856 GCATCAGAGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGATTTGT 2915
Db      1785 GCATCAGAGAGATGTCAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGATTTGT 1844
QY      2916 GTACAGCATGCGCTTCACAAATGAGATGATGA 2949
Db      1845 GTACAGCATGCGCTTCACAAATGAGATGATGA 1878

```

RESULT 3
ABX81661
ID ABX81661 standard; cDNA; 272 BP.

XX ABX81661;
AC
XX
XX
DT 24-APR-2003 (first entry)

DE Corn ear-derived polynucleotide (cpd) #121.

XX Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
XX structural gene; functional gene; regulatory gene;
XX corn ear-specific profile; gene transcription; gene expression;
XX hybrid plant; desirable trait expression; plant breeding program;
XX inheritance; desired characteristic; growth; development;
XX disease resistance; environmental adaptability; quality; yield;
XX multigene trait; plant; gene; ss.

XX Zea mays.

OS US6476212-B1.

XX 05-NOV-2002.

XX 14-MAY-1999; 99US-00313294.

XX 26-MAY-1998; 98US-0086722P.

XX (INCY-) INCYTE GENOMICS INC.

XX LaJuvdi RV, Ito LY, Sherman BK;

XX WPI; 2003-208840/20.

XX Novel purified corn-ear derived polynucleotide useful as hybridization
XX probe for detecting polynucleotide in sample, and for identifying,
XX evaluating, and altering desired characteristics associated with growth,
XX development.

XX Example; SEQ ID NO 121; 390pp; English.

XX The present invention relates to the isolation of corn ear-derived
XX polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
XX and SATMON023. Some of the cdps uniquely identify structural, functional,
XX and regulatory genes of corn ear. The polynucleotide sequences are
XX useful for detecting cdps in a sample, for producing a corn ear-specific
XX profile of gene transcription, for detecting altered gene expression in
XX inbred or hybrid plants, and for screening several molecules for specific
XX binding to the polynucleotide. The cdps are useful to identify, isolate,

CC or extend identical or related corn-ear nucleic acid sequences from DNA
CC libraries, and in nucleic acid amplification or hybridization techniques
CC to follow the expression of desirable traits through plant breeding
CC programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC follow the inheritance of desired characteristics associated with growth
CC and development, disease resistance, environmental adaptability, quality,
CC and yield of corn. The cdps are also useful as molecular markers for
CC studying inheritance and multigene traits in a plant breeding program.
CC The cdps are useful for producing purified corn-ear polypeptides by
CC recombinant techniques. They are also useful in diagnostic assays to
CC detect or confirm conditions or diseases associated with abnormal levels
CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC polynucleotides (cdps) of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/patseq/identity.html

SQ Sequence 272 BP; 79 A; 52 C; 68 G; 67 T; 0 U; 6 Other;

Query Match 2.3%; Score 74.4; DB 7; Length 272;

Best Local Similarity 58.7%; Pred. No. 7.8e-08;

Matches 142; Conservative 0; Mismatches 97; Indels 3; Gaps 1;

QY 1777 TATGAGCTCTTAGAGACGCCATTCTATGTCACAGGCGCATGAGATGAGTGT 1836

Db 2 TATGAGCAAGTMAAGCAGTCATGCTATNGCTCCCTGCGCACCGTGTATGAGCGTN 61

QY 1837 CTGATGTTTGAAGACAGTCGCCACTGGCTATTTGAGAGCGCGAAACGCTCCACCGGAGTTA 1896

Db 62 TTAATATTTGAAAGCTCAGCTGCTGAGCTGACATGGAAGCTGGAAGCTGTGATTAACACTTT 121

QY 1897 GCGTGAATGGGTTGATGATGATGCTGG---GCTCAGAACGCGATATGTTTCTGGA 1953

Db 122 GTTAAACAAAGTACAGACAGAAATTCATGCACTTACGCAAGTTGATTTGGCTGAT 181

QY 1954 GGTGTTGCGCAACTGATGCTTCTTCCAGCAAGCAAGATCTGACATATTCATCAA 2013

Db 182 GGGAAAGCCNCCTATATGTTCTTACCAAAAGAGATATGAGACCATTTTAAACAAG 241

QY 2014 CA 2015

Db 242 CA 243

XX

RESULT 4
ABX82003
ID ABX82003 standard; cDNA; 274 BP.

XX ABX82003;

XX 24-APR-2003 (first entry)

XX Corn ear-derived polynucleotide (cpd) #463.

XX Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
XX structural gene; functional gene; regulatory gene;
XX corn ear-specific profile; gene transcription; gene expression;
XX hybrid plant; desirable trait expression; plant breeding program;
XX inheritance; desired characteristic; growth; development;
XX disease resistance; environmental adaptability; quality; yield;
XX multigene trait; plant; gene; ss.

XX Zea mays.

OS US6476212-B1.

XX 05-NOV-2002.

XX 14-MAY-1999; 99US-00313294.

XX 26-MAY-1998; 98US-0086722P.

XX (INCY-) INCYTE GENOMICS INC.

XX Lalgudi RV, Ito LY, Sherman BK;
 XX WPI; 2003-208840/20.
 XX Novel purified corn-ear derived polynucleotide useful as hybridization
 PT probe for detecting polynucleotide in sample, and for identifying,
 PT evaluating, and altering desired characteristics associated with growth,
 PT development.

PS Example; SEQ ID NO 463; 390bp; English.

XX The present invention relates to the isolation of corn ear-derived
 CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
 CC and SATMON023. Some of the cdps uniquely identify structural, functional,
 CC and regulatory genes of corn ear. The polynucleotide sequences are
 CC useful for detecting cdps in a sample, for producing a corn ear-specific
 CC profile of gene transcription, for detecting altered gene expression in
 CC inbred or hybrid plants, and for screening several molecules for specific
 CC binding to the polynucleotide. The cdps are useful to identify, isolate,
 CC or extend identical or related corn-ear nucleic acid sequences from DNA
 CC libraries, and in nucleic acid amplification or hybridisation techniques
 CC to follow the expression of desirable traits through plant breeding
 CC programs. Preferably, the cdps are used to identify, evaluate, alter, or
 CC follow the inheritance of desired characteristics associated with growth,
 CC and development, disease resistance, environmental adaptability, quality,
 CC and yield of corn. The cdps are also useful as molecular markers for
 CC studying inheritance and multigene traits in a plant breeding program.
 CC The cdps are useful for producing purified corn-ear polypeptides by
 CC recombinant techniques. They are also useful in diagnostic assays to
 CC detect or confirm conditions or diseases associated with abnormal levels
 CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
 CC polynucleotides (cdps) of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/psipdsIDentry.html

XX Sequence 274 BP; 83 A; 49 C; 73 G; 67 T; 0 U; 2 Other;

XX Query Match 2.2%; Score 72.2; DB 7; Length 274;

XX Best Local Similarity 60.6%; Pred. No. 2.7e-07; Matches 152; Conservative 0; Mismatches 95; Indels 4; Gaps 2;

QY 1777 TATGAGGCTTATGAGACGCGCATTCCTATGAT-CCACAGGCGCATCGGATGAGTGT 1835
 Db |||||
 QY 2 TATGAGGAGCAATGATGAGAGCTGATGCTATGCTCTCGGGCACCGTGAGATGAGCT 61
 Db |||||
 QY 1836 TCTGATGTTTGAAGACAGTCCACTGCTATTTTGAAGGCCGAAAGCCTCCACCGGAGTT 1895
 Db |||||
 QY 62 GTTAAATATTTGAAAGCTCAGCTGTGGCTACATGAGGAAGCTGAACGTCCTCATMAACATT 121
 Db |||||
 QY 1896 AGCTGAGATGGGTTAGATGATGCTGCG--GGTCAAGAACGAGATGATGTTTCTG 1952
 Db |||||
 QY 122 TGTTAATCAAGATGACAGACGAAATTCATGACCTAACGCAAGGTTGATTTGCTG 181
 Db |||||
 QY 1953 AGGTGTTCCGCAACTGATGAGCTTCTTGCAAGAAAGCAAGATCTGACATATTCATCA 2012
 Db |||||
 QY 182 TGGGAAAGGCACTATGATGTTCTTAGCAACAAAGAGATGAGAGCAATTAAACA 241
 Db |||||
 QY 2013 AACCTCTCAAG 2023
 Db |||||
 QY 242 GCATTGCCAGG 252
 Db |||||

RESULT 5

ABL33697

ID ABL33697 standard; DNA; 6668 BP.

AC ABL33697;

DT 26-MAR-2002 (first entry)

DE Human immune system associated gene SEQ ID NO: 1670.

XX Human; immune system disease; cytosine methylation; antiaesthetic;
 KW antiarteriosclerotic; antiamebic; cytosolic; noctropic;
 KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
 KW antihemematic; antiarthritic; antidiabetic; antipsoriatic;
 KW antineuritic; cancer; eye disease; arteriosclerosis; anaemia;
 KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
 KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
 KW de.

XX Homo sapiens.

XX WO200200928-A2.

XX 03-JUN-2002.

XX 02-JUL-2001; 2001WO-BP007537.

XX 30-JUN-2000; 2000DE-01032529.

XX 01-SEP-2000; 2000DE-01043826.

XX (EPIC-) EPIGENOMICS AG.

XX Olek A, Piepenbrock C, Berlin K;

XX WPI; 2002-130909/17.

XX Nucleic acid comprising fragment of chemically modified gene, useful for
 PT diagnosis and treatment of diseases associated with abnormal cytosine
 PT methylation.

PS Claim 1; SEQ ID NO 1670; 322p + Sequence Listing; German.

XX The present invention provides a number of human immune system associated
 CC genes which are modified by the methylation of cytosines. The sequences
 CC can be used in the diagnosis and treatment of immune system disorders,
 CC including eye diseases such as retinopathy, neovascular glaucoma and
 CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
 CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
 CC rheumatoid arthritis, psoriasis and inflammatory/alicerative bowel
 CC diseases. The present sequence is a gene of the invention

XX Sequence 6668 BP; 1329 A; 328 C; 1736 G; 3274 T; 0 U; 1 Other;

XX Query Match 2.1%; Score 68.8; DB 6; Length 6668;

XX Best Local Similarity 45.8%; Pred. No. 7.6e-06; Matches 238; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

QY 156 GTTTAGAGCTTAATTAAGCTTCCTCATTTGCTCTTCTGTCAGTTATTTCTCTC 215
 Db |||||
 QY 2742 GTTTTGTGTTTGT 2801
 Db |||||
 QY 216 CGAAGTCTGACTACATCTCTCGGCGCTTAACTTACGTTCTCGTCTT 275
 Db |||||
 QY 2802 TT 2861
 Db |||||
 QY 276 ACTGTAAGTTTCTGCTAGAGCCCGCATGCGCACCAGATGATCTGCTCG 335
 Db |||||
 QY 2862 TTTTTTGTGTTT 2921
 Db |||||
 QY 336 ATTCTCTTTTCTGCTGGAATAATGCGCTAATGTTCTGATTTGAGAGTTTGT 395
 Db |||||
 QY 2922 TTTTGTGTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 2981
 Db |||||
 QY 396 GCTATGAGTTATTTTCCCTAATATTTATATGTTCTTAAAGTAACATCTGCTCTTA 455
 Db |||||
 QY 2982 TTTTGTGTTT 3041
 Db |||||
 QY 456 CTGTTTGTTCATTTTGTGCTTTCACGTTTAAAGCGTATCGAGATTTGACTG 515
 Db |||||
 QY 3042 TTTTGTGTTT 3101
 Db |||||
 QY 516 TGAATAATCCTTCGTTTTTGTGTTTTCATATAAATGAGATGATCTACCTTTGT 575
 Db |||||

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

Query Match 2.1%; Score 68.6; DB 5; Length 556;
 Best Local Similarity 48.3%; Pred. No. 2.8e-06;
 Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

265 CTCGGTGGTCTGCTGTAAGTTTCTGCTTGAAGCCCTCGATCGCTCAGCGCATGCA 324
 486 CCCCCCTCCCCCTTTTAAATATTTTATTTTCTCCGCCCTCCCATTAATA 427
 325 TTCGTCGCTGATTCCTTTCTTTCGCTGGAATAATGCCCTAATGTCGATTCG 384
 426 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 367
 385 AAGGTTTGTGCTAGGTTACTTTTCCCTATATTTATGTTCTTAGTAAGATA 444
 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
 445 CTGCGCTTACTGTTTGTCTATTTGTTGCTTTCACCGCTTACGCTGATCGGA 504
 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
 505 GTATTGACCTGGAATAATCCTCGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 564
 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
 565 CTACCTTTGTGCTTGAATGTTGTTTGTGAGCCTAAGCGTTGCTGTATTAAGT 624
 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
 625 TCACGTTCAATGCTGATTTTGAGATTTT 653
 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98

RESULT 8

ABV42105/c
 ID ABV42105 standard; cDNA; 556 BP.
 XX
 AC ABV42105;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 42096.
 XX
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KM pharmacogenomic marker; gene; ss.
 KM
 OS Homo sapiens.
 OS
 PN WO200160860-A2.
 PN
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001MO-US005171.
 XX
 XX 17-FEB-2000; 2000US-018319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 PR
 XX
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PA
 PA Schlegel R, Endege WO, Monahan DE;
 PI

XX
 DR WPI; 2001-662795/76.

XX
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer; useful
 PT for detecting presence of prostate cancer; stage of prostate cancer.
 XX
 PS
 XX Claim 1; Page 8441; 11750bp; English.

XX The invention relates to an isolated nucleic acid molecule (i) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (i) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

Query Match 2.1%; Score 68.6; DB 5; Length 556;
 Best Local Similarity 48.3%; Pred. No. 2.8e-06;
 Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

265 CTCGGTGGTCTGCTGTAAGTTTCTGCTTGAAGCCCTCGATCGCTCAGCGCATGCA 324
 486 CCCCCCTCCCCCTTTTAAATATTTTATTTTCTCCGCCCTCCCATTAATA 427
 325 TTCGTCGCTGATTCCTTTCTTTCGCTGGAATAATGCCCTAATGTCGATTCG 384
 426 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 367
 385 AAGGTTTGTGCTAGGTTACTTTTCCCTAATTTATGTTCTTAGTAAGATA 444
 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
 445 CTGCGCTTACTGTTTGTCTATTTGTTGCTTTCACCGCTTACGCTGATCGGA 504
 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
 505 GTATTGACCTGGAATAATCCTCGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 564
 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
 565 CTACCTTTGTGCTTGAATGTTGTTTGTGAGCCTAAGCGTTGCTGTATTAAGT 624
 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
 625 TCACGTTCAATGCTGATTTTGAGATTTT 653
 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98

RESULT 9

ABV43601/c
 ID ABV43601 standard; cDNA; 556 BP.
 XX
 AC ABV43601;
 XX
 DT 16-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 43592.
 XX
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KM pharmacogenomic marker; gene; ss.
 KM
 OS Homo sapiens.
 OS
 PN WO200160860-A2.
 PN

XX 23-AUG-2001.
 PD 20-FEB-2001; 2001WO-US005171.
 XX 17-FEB-2000; 2000US-0183319P.
 XX 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 PI Schlegel R, Endege WO, Monahan JE.
 XX WPI; 2001-662795/76.
 DR Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 8681-8682; 117500P; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 CC XX
 SQ Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;
 Query Match 2.1%; Score 68.6; DB 5; Length 556;
 Best Local Similarity 48.3%; Pred. No. 2.8e-06;
 Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;
 QY 265 CTCGCGCTTACTGCTGAGTTTTCGCTTACGCTCGGCTGCGCTGCGGATGCA 324
 DB 486 CMCCTCTCCCTTTTATTTATTTTATTTTCTCCCTCCCTCCCATTAATA 427
 QY 325 TTCTGCTCGATTTCTCTTTCGCTGGAATAATGCTTAAATGCTCGATTTGCG 384
 DB 426 TTTTCTTTTCTTTTCTTTTCTTTTATAAATAAATACNCCCTTTTCTTTT 367
 QY 385 AAGCTTTTGTGCTAGGCTTCTTTCCTTAATTTTAAATGTTCTTAAAGTAAGA 444
 DB 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
 QY 445 CCGGCTTACTGCTTGTGCTTGTGCTTACCGTTAGTGCATGCGA 504
 DB 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
 QY 505 GATTTGACTGTGAAATCCCTGTTTGTGTTTGTGCTATTAATGCGATGAT 564
 DB 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
 QY 565 CTACCTTTGTGCTTGTGATTTGTTTGTGACCTTAAAGGTTGCTTAACT 624
 DB 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
 QY 625 TCAGCTTCAATGATTTTGAATTT 653
 DB 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98

AA545347
 ID AA545347 standard; DNA; 9539 BP.
 XX
 AC AA545347;
 XX
 DT 18-DEC-2001 (first entry)
 XX
 DE Chemically pretreated complementary DNA associated with cell cycle #26.
 XX
 KW Cell cycle; human; CpG dinucleotide; cytosine methylation; HIV; aging;
 KW human immunodeficiency virus; neurodegenerative disorder; solid tumour;
 KW graft-versus-host disease; glomerular disease; Lewy body disease; cancer;
 KW arthritis; arteriosclerosis; anti-HIV; neuroprotective; antiarthritic;
 KW immunosuppressive; antitumour; cytostatic; antiarteriosclerotic; ds;
 KW PCR primer.
 XX
 OS Homo sapiens.
 XX
 PN WO200168911-A2.
 XX
 PD 20-SEP-2001.
 XX
 PF 15-MAR-2001; 2001WO-EP002945.
 XX
 PR 15-MAR-2000; 2000DE-01013847.
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2001-602751/68.
 XX
 PT Designing primers and probes for analyzing diseases associated with
 PT cytosine methylation state e.g. arthritis, cancer, aging,
 PT arteriosclerosis comprising fragments of chemically modified genes
 PT associated with cell cycle.
 XX
 PS
 XX
 SQ Claim 1; SEQ ID NO 52; 28pp; English.
 CC Sequences AA545296-AA545520 represent chemically pretreated genomic DNA
 CC molecules associated with the cell cycle and specific PCR primers of the
 CC invention. The sequences are useful for detecting the methylation state
 CC of all CpG dinucleotides in a sequence and therefore for analysing
 CC associated diseases. By analysing cytosine methylations in the pretreated
 CC DNA, genetic and/or epigenetic parameters for the diagnosis and therapy
 CC of existing diseases or the predisposition to specific diseases can be
 CC ascertained. The parameters may be compared to another set of genetic
 CC and/or epigenetic parameters, the differences serving as basis for
 CC diagnosis and/or prognosis events which are disadvantageous to patients.
 CC The sequences of the invention are useful for the diagnosis and therapy
 CC of HIV infection, neurodegenerative disorders, graft-versus-host disease,
 CC aging, glomerular disease, Lewy body disease, arthritis,
 CC arteriosclerosis, solid tumours and cancers
 CC XX
 SQ Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 U; 0 Other;
 Query Match 2.0%; Score 66.8; DB 4; Length 9539;
 Best Local Similarity 46.7%; Pred. No. 2.7e-05;
 Matches 212; Conservative 0; Mismatches 242; Indels 0; Gaps 0;
 QY 251 TTAAACTTACGTTCTCGCTGTTTACCTGTAAGTTTCTGCTTAAAGCCTCGATCG 310
 DB 202 TTTTCTTTTGTATTTTTCGTTTATTTTGTGCTTTTCTTTTCTTTTCTTTTCTTTT 261
 QY 311 CCGACCGGATGATCTGTCGATTTCTCTTTTCTGCTGGAATAATGCGCTTA 370
 DB 262 GTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 321
 QY 371 TGTCTCGATTTGGAAGTTTGTGCTATGAGTTATTTTCCCTATATTTTAAAGTT 430

KW Nijmegen breakage syndrome; Werner syndrome; immunodeficiency;
 KW trichothiodystrophy; Fanconi's anaemia; solid tumour; cancer; ds.
 XX
 OS Unidentified.
 XX
 PN WO200181622-A2.
 XX
 PD 01-NOV-2001.
 XX
 PF 06-APR-2001; 2001MO-EP003972.
 XX
 PR 06-APR-2000; 2000DE-01019058.
 PR 07-APR-2000; 2000DE-01019173.
 PR 30-JUN-2000; 2000DE-01032529.
 PR 01-SEP-2000; 2000DE-01043826.
 XX
 PA (EPIG-) EPIGENOMICS AG.
 XX
 PI Olek A, Piepenbrock C, Berlin K;
 XX
 DR WPI; 2002-034446/04.
 XX
 PT New nucleic acid derived from genes associated with DNA repair, useful
 PT for diagnosis, e.g. of ataxia telangiectasia, by determination of
 PT cytosine methylation.
 XX
 PS Claim 1; SEQ ID NO i22; 25pp + Sequence Listing; English.
 CC The invention relates to nucleic acids containing a sequence of at least
 CC 18 nucleotides of chemically treated DNA of genes associated with DNA
 CC repair, and their complements. The invention also relates to nucleic
 CC acids comprising at least 18 base pairs of the chemically pretreated DNA
 CC of genes associated with DNA repair selected from PMS2L1, PMS2L2,
 CC PMS2L3, PMS2, L4, PMS2L5, MGMT, MSH2, NUDT11, TDG, INPPL1,
 CC RFC4, DIRT1L, FANCB, or XRCC8. Nucleic acids of the invention and related
 CC oligomers, are useful for diagnosis of diseases associated with gene
 CC repair, specifically ataxia telangiectasia, aging, Bloom's syndrome,
 CC Cockayne syndrome, Nijmegen breakage syndrome or Werner syndrome.
 CC immunodeficiency, trichothiodystrophy, Fanconi's anaemia, solid tumours
 CC and cancer, particularly by determining status of cytosine methylation
 CC and/or by detecting single-nucleotide polymorphisms. Determination of
 CC individual methylation patterns may allow development of individualised
 CC therapies. The sequences given in records ABR92192-ABR92335 represent
 CC chemically pre-treated DNA fragments from genes associated with DNA
 CC repair, and their complements. Note: The sequence data for this patent is
 CC not represented in the specification, but is based on sequence
 CC information supplied by the European Patent Office
 CC
 SQ Sequence 8079 BP; 1992 A; 212 C; 1962 G; 3913 T; 0 U; 0 Other;
 Query Match 2.0%; Score 66.4; DB 6; Length 8079;
 Best Local Similarity 48.2%; Pred. No. 3.2e-05;
 Matches 187; Conservative 0; Mismatches 201; Indels 0; Gaps 0;
 QY 333 TCGATTTCTTTTCTTCGCGTGAAGAAATGCGCTAATGTTCCATTTGCAAGGTTT 392
 DB 5327 TGTGTTATGTTTATATATTAGTGAAGCGCTTATGTTGATTTTGTCTTTT 5386
 QY 393 TGTGCTAGGGTACTTTTCCCTAATTTTAAAGTTCTTAGTAACGATACCTGCTC 452
 DB 5387 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5446
 QY 453 TTAGCTTTTGTTCATTTTGTGTCCTTACCGTTTACGCTGATCGAGATTGTA 512
 DB 5447 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5506
 QY 513 CAGTGAAGAAATCCTTGTTTGTGTTTGTTCATTAATTAATCGATGATCTTACCTTT 572
 DB 5507 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5566
 QY 573 TGTGCTTGAAGTTTCTTTTGTGAGCCATGCGTGTGGCTGTATAACTTCACGCTT 632
 DB 5567 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5626

QY 633 ATGTGTGATTTTGAAGATTTTGTGATGACGTGGGTTTCTTGTGCTATAGCTTGA 692
 DB 5627 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 5686
 QY 693 AAAATGAGTTCTAGGCGTGTGCAATGT 720
 DB 5687 GTGTAGTGTATGATTATGATTATGT 5714
 RESULT 13
 AAX33181/C
 ID AAX33181 standard; DNA; 6644 BP.
 XX
 AC AAX33181;
 XX
 DT 25-JUN-1999 (first entry)
 XX
 DE Base sequence of the plasmid pRX-ires-bar.
 XX
 KW Compo virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
 KW bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer;
 KW autoimmune disease; graft rejection reaction; inflammation;
 KW inflammatory disease; ss.
 XX
 OS Synthetic.
 OS Compo virus.
 XX
 PN WO9913073-A2.
 XX
 PD 18-MAR-1999.
 XX
 PF 07-SEP-1998; 98WO-UP004010.
 XX
 PR 08-SEP-1997; 97JP-00259235.
 XX
 PA (RPRG-) RPR GENCELL ASIA PACIFIC INC.
 XX
 PI Hamada H;
 XX
 DR WPI; 1999-243728/20.
 XX
 PT New apoptosis-resistant virus-sensitive cell.
 XX
 PS Example 1; Page 38-41; 51pp; English.
 XX
 CC The present invention describes an apoptosis-resistant virus-sensitive
 CC cell line into which an apoptosis resistance gene has been introduced.
 CC The recombinant viruses generated are capable of expressing apoptosis-
 CC associated genes. These can then be used in a variety of diseases for
 CC which the induction of apoptosis by gene transfer, or where the
 CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
 CC are useful as vectors for gene therapy which can be applied to cancer
 CC therapy for destroying cancer cells selectively, the treatment of
 CC autoimmune diseases and graft rejection reaction, and apoptosis induction
 CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
 CC encountered the problem where if an adenovirus vector capable of
 CC expressing an apoptosis-associated gene is introduced into animal cells,
 CC the cells producing the virus will be destroyed because the period of
 CC time required to induce cell death by apoptosis is shorter than that
 CC required to replicate and produce the virus, resulting in failure to
 CC obtain a recombinant virus having the integrated apoptosis-associated
 CC gene. In this invention an apoptosis-resistant 293 cell line (having an
 CC apoptosis resistant gene introduced) is established and overcomes the
 CC problem. The present sequence represents the base sequence of the plasmid
 CC pRX-ires-bar, which contains the compo virus bsr gene, and is used in an
 CC example from the present invention
 CC
 SQ Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 U; 0 Other;
 Query Match 2.0%; Score 66; DB 2; Length 6644;
 Best Local Similarity 45.4%; Pred. No. 3.6e-05;
 Matches 237; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 14:30:53 ; Search time 8548.15 Seconds
(without alignments)
16605.743 Million cell updates/sec

Title: US-10-030-829-1
Perfect score: 3275
Sequence: 1 gacaaacaacaaatctaa.....caaatatagggtacaaat 3275

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank1:
1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rtd:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcg_hum:*
40: em_hcg_mus:*
41: em_hcg_other:*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3275	100.0	3275	6	AX078760
2	3275	100.0	81365	8	AB025633
3	2254	68.8	2254	8	AF239719
4	1651.2	50.4	2162	8	BT002944
5	1493	45.6	1909	8	BT004380
6	1462	44.6	1878	6	AX078761
7	643.6	19.7	650	8	ATHS28171
8	194	5.9	1556	8	AK064217
9	189.8	5.8	2181	8	AK064995
10	181.8	5.6	94911	8	ATT14E10
11	174	5.3	91053	8	CNS08CNO
12	174	5.3	112231	8	CNS07Y01
13	155.8	4.8	2570	8	AK100659
14	120.8	3.7	1564	8	AF542974
15	117	3.6	513	8	AF469493
16	112.8	3.4	1214	8	AK061975
17	85.6	2.6	7218	6	I66494
18	83.4	2.5	184535	6	CNS05TCJ
19	81.4	2.5	349980	6	AX344566
20	74.4	2.3	272	6	AR244762
21	72.8	2.2	1407	8	AJ592026
22	72.2	2.2	274	6	AR245104
23	71	2.2	60565	2	AC023852
24	70	2.1	1453	8	AJ591978
25	70	2.1	66993	2	AC138074
26	69.6	2.1	1434	8	AJ592058
27	69.2	2.1	3364	14	TSP418778
28	69	2.1	88932	2	AC022680
29	68.8	2.1	6668	6	AX346599
30	68.8	2.1	81120	2	AC022851
31	68.6	2.1	7218	6	I66494
32	68.4	2.1	810	8	AJ592180
33	68.2	2.1	56857	2	AC021917
34	68.2	2.1	194180	2	AC069140
35	68	2.1	1184	8	AJ591982
36	67.6	2.1	151368	2	AC142664
37	67.4	2.1	156550	2	AC015830
38	67.4	2.1	196853	9	CNS01PUD
39	67.4	2.1	273807	2	AC025421
40	67	2.0	1228	8	AJ592059
41	67	2.0	127709	2	CNS05TC1
42	67	2.0	202083	2	AC023833
43	66.8	2.0	9539	6	AX277889
44	66.8	2.0	9539	6	AX323566
45	66.8	2.0	70389	2	AC135853

ALIGNMENTS

RESULT 1
AX078760 3275 bp DNA linear PAT 22-FEB-2001
LOCUS Sequence 1 from Patent WO0105951.
DEFINITION
ACCESSION AX078760
VERSION AX078760.1 GI:13158379
KEYWORDS
SOURCE
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Beclin,C., Eimayan,T. and Vaucherec,H.
TITLE Novel sgsl plant gene and use thereof

Pred. No. is the number of results predicted by chance to have a

JOURNAL

Patent: WO 0105951-A 1 25-JAN-2001;
 AVENTIS CORPSCIENCE S.A. (FR) ; INSTITUT NATIONAL DE LA RECHERCHE
 AGRONOMIQUE (FR)

FEATURES

Location/Qualifiers

source 1..3275

/organism="Arabidopsis thaliana"

/mol_type="unassigned DNA"

/db_xref="taxon:3702"

693..715

/note="p356AD"

complement(2926..2952)

/note="p356Y"

ORIGIN

Query Match 100.0%; Score 3275; DB 6; Length 3275;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 GACAAACAAACAAATTAAGCAAGTCATGTCGTCGATTAATTAATTAAGTGGACAA 60
 1 GACAAACAAACAAATTAAGCAAGTCATGTCGTCGATTAATTAATTAAGTGGACAA 60
 61 TTAAGTTAAGCGAAAAAGAAAAAGGTACAAAAATGAAAAACAAATCAACTGAAT 120
 61 TTAAGTTAAGCGAAAAAGAAAAAGGTACAAAAATGAAAAACAAATCAACTGAAT 120
 121 GAAATTTGAGTCAGAAATCGAAAAACGAGCCGTTTGAAGTTAATTAAGTCTCTC 180
 121 GAAATTTGAGTCAGAAATCGAAAAACGAGCCGTTTGAAGTTAATTAAGTCTCTC 180
 181 ATTTGTCCTCTTCTGTCAGTTATTTCTCTCCGAGTCCTGACTCACTCTCAC 240
 181 ATTTGTCCTCTTCTGTCAGTTATTTCTCTCCGAGTCCTGACTCACTCTCAC 240
 241 TCTCCGAGGCTTTAACTTAAGTCTCTCCGTCGTTACTGTAAGTTTCTGCTTAG 300
 241 TCTCCGAGGCTTTAACTTAAGTCTCTCCGTCGTTACTGTAAGTTTCTGCTTAG 300
 301 CCTCCGATCGCCTACCGCATGATCTGTCGATTTCTTTTCTTCTGCTGAAAA 360
 301 CCTCCGATCGCCTACCGCATGATCTGTCGATTTCTTTTCTTCTGCTGAAAA 360
 361 ATTCGCCCTAATGTCGATTCGAAAGTTTGTGCTAAGGTTACTTTTCCCTATA 420
 361 ATTCGCCCTAATGTCGATTCGAAAGTTTGTGCTAAGGTTACTTTTCCCTATA 420
 421 TTTTAAAGTCTTAAGTAACGATACCTGCGTCTTAAGTTTGTCAATTTGTTGCT 480
 421 TTTTAAAGTCTTAAGTAACGATACCTGCGTCTTAAGTTTGTCAATTTGTTGCT 480
 481 TTCACCGTTTGTGCTGATCGAGTATTGACGTGAAAAATCCTCGTTTGGTTT 540
 481 TTCACCGTTTGTGCTGATCGAGTATTGACGTGAAAAATCCTCGTTTGGTTT 540
 541 TTGTTCAATTAATCGAATGATCTACCTTTTGTGCTTGTGATGTTTGTGAGCT 600
 541 TTGTTCAATTAATCGAATGATCTACCTTTTGTGCTTGTGATGTTTGTGAGCT 600
 601 ATGCGTTGTGCTGTTATAAATTACGTTCAATGTCGATTTGAATTTGGTAGTG 660
 601 ATGCGTTGTGCTGTTATAAATTACGTTCAATGTCGATTTGAATTTGGTAGTG 660
 661 ACTGTCGTTTCTTTGTCGCTAATGTTGTAATAAGTCTAAGGCTGATCCAAATG 720
 661 ACTGTCGTTTCTTTGTCGCTAATGTTGTAATAAGTCTAAGGCTGATCCAAATG 720
 721 CTAAGAAAAAAGACGTTCAAGGCTGTTAAGCCTGAAGTTGAACAGTTGTTCAAGTT 780
 721 CTAAGAAAAAAGACGTTCAAGGCTGTTAAGCCTGAAGTTGAACAGTTGTTCAAGTT 780
 781 TGGCAGGAGCAGAGACTGCTTCTTCAAGATGATGAGAGAGAGTGGAGGTCATTTTCA 840
 781 TGGCAGGAGCAGAGACTGCTTCTTCAAGATGATGAGAGAGAGTGGAGGTCATTTTCA 840

841 AGAAGAACAAAGAACAAACCGAAGACCTTGGAAAACTTGGGTTTCTCAGAAATTCGA 900
 841 AGAAGAACAAAGAACAAACCGAAGACCTTGGAAAACTTGGGTTTCTCAGAAATTCGA 900
 901 ATCTCTTAGAGCTTGGGGTGTGTCAGCAGCAAGGAGAGTACCAAGTATCGGAGAG 960
 901 ATCTCTTAGAGCTTGGGGTGTGTCAGCAGCAAGGAGAGTACCAAGTATCGGAGAG 960
 961 GAAACAAATGATCCGGAGAGGTAACCGCAATGTCGGGGCATTAAGGCTAACATATCTG 1020
 961 GAAACAAATGATCCGGAGAGGTAACCGCAATGTCGGGGCATTAAGGCTAACATATCTG 1020
 1021 GTGGGGACGAGCGTTGAGCAGAAAGTATGATTAACAATTTGTGGACCCCACTGTAT 1080
 1021 GTGGGGACGAGCGTTGAGCAGAAAGTATGATTAACAATTTGTGGACCCCACTGTAT 1080
 1081 CTCGCCCTCTTTGAAAGAGATGAAATTTGGCAGGCAAGAGAGTCTTCTGAGGAGAGA 1140
 1081 CTCGCCCTCTTTGAAAGAGATGAAATTTGGCAGGCAAGAGAGTCTTCTGAGGAGAGA 1140
 1141 CAGCTGTCAGAGATTTCTGACGTGAGATGATGATGATGATGATGATGATGATGATG 1200
 1141 CAGCTGTCAGAGATTTCTGACGTGAGATGATGATGATGATGATGATGATGATGATG 1200
 1201 ATGATTCGATGCTTTGGATGATTTCTGATGACGACCTTGCAGATGATGATGATGATG 1260
 1201 ATGATTCGATGCTTTGGATGATTTCTGATGACGACCTTGCAGATGATGATGATGATG 1260
 1261 ATGATGATCAAAAGAGCCATGATGATCAAGAAAGCAAAATTAAGTGTTCAAAAAGTCTT 1320
 1261 ATGATGATCAAAAGAGCCATGATGATCAAGAAAGCAAAATTAAGTGTTCAAAAAGTCTT 1320
 1321 GAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
 1321 GAGCTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1380
 1381 CAGCTTGTCAAGACGAGCTGTCGATGATGATGATGATGATGATGATGATGATGATG 1440
 1381 CAGCTTGTCAAGACGAGCTGTCGATGATGATGATGATGATGATGATGATGATGATG 1440
 1441 ATGCGAGACCAAAAGAGCTGAGGCAATGATGATGATGATGATGATGATGATGATGATG 1500
 1441 ATGCGAGACCAAAAGAGCTGAGGCAATGATGATGATGATGATGATGATGATGATGATG 1500
 1501 AAAAGATCTACAGATGAGAGCGCATCTGTCATTTCTTGTGATGATTTATGGGCAAT 1560
 1501 AAAAGATCTACAGATGAGAGCGCATCTGTCATTTCTTGTGATGATTTATGGGCAAT 1560
 1561 GGAAGGTTGGGTGAGATGAAAAAGATTAAGAAATTTGCTGCTCAATGTCATCA 1620
 1561 GGAAGGTTGGGTGAGATGAAAAAGATTAAGAAATTTGCTGCTCAATGTCATCA 1620
 1621 TCAATGAATCTAGACTGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 1680
 1621 TCAATGAATCTAGACTGATTAAGAGATTAAGAGATTAAGAGATTAAGAGATTAAGAG 1680
 1681 TTCTTTAATTTTCTTGCATTTCTAATGATCTTAAGATTTGAATTTGAATTTGAATTT 1740
 1681 TTCTTTAATTTTCTTGCATTTCTAATGATCTTAAGATTTGAATTTGAATTTGAATTT 1740
 1741 ATGGGCAACCAAGAGCTGCTGAAATCTTGAACAGTATGAGAGCTTTAGACAGCCAT 1800
 1741 ATGGGCAACCAAGAGCTGCTGAAATCTTGAACAGTATGAGAGCTTTAGACAGCCAT 1800
 1801 TCCATGATCCACAGGCGCATGTCGATGATGATGATGATGATGATGATGATGATGATG 1860
 1801 TCCATGATCCACAGGCGCATGTCGATGATGATGATGATGATGATGATGATGATGATG 1860
 1861 GGCATTTGAGAGCCGAAAGCCTCTCAACCGGAGATTAAGTGAATTTAGTAAATTT 1920
 1861 GGCATTTGAGAGCCGAAAGCCTCTCAACCGGAGATTAAGTGAATTTAGTAAATTT 1920

Oy	1921	GCCTGGGGCTCAAGACGAGATATGTTTTCTGAGAGTGTGGCCAACTGATAGGGCTCTT	1990
Dp	1921	GCCTGGGGCTCAAGACGAGATATGTTTTCTGAGAGTGTGGCCAACTGATAGGGCTCTT	1980
Oy	1981	GCAACGAAGCAAGATCTGAGACATATTCATCAAGACTCTCAAGGTTCTCTCCCCAAGA	2040
Dp	1981	GCAACGAAGCAAGATCTGAGACATATTCATCAAGACTCTCAAGGTTCTCTCCCCAAGA	2040
Oy	2041	AATTTGATATATGCTTTTATGTTTGTCAATGGAATTTAAAGTTTGTGTGCGTGTAA	2100
Dp	2041	AATTTGATATATGCTTTTATGTTTGTCAATGGAATTTAAAGTTTGTGTGCGTGTAA	2100
Oy	2101	TGCATCTGTATATATATATCTATATGATTCATTAGCCAAACAAAGGCTGAATTGAGTTG	2160
Dp	2101	TGCATCTGTATATATATATCTATATGATTCATTAGCCAAACAAAGGCTGAATTGAGTTG	2160
Oy	2161	AAATCATATCCAAAGAGATGTTGTAAAGAGCTGAGGCGAGATCTCTGAGAGCAATACAGCAG	2220
Dp	2161	AAATCATATCCAAAGAGATGTTGTAAAGAGCTGAGGCGAGATCTCTGAGAGCAATACAGCAG	2220
Oy	2221	CTGAACCTACTTTTAAAGAACAACTCTCAAAACAGAACAGACCGCAAGTCTTGAGGAA	2280
Dp	2221	CTGAACCTACTTTTAAAGAACAACTCTCAAAACAGAACAGACCGCAAGTCTTGAGGAA	2280
Oy	2281	TCTCTGGAATAATTAGAGCGAGAAAGTGGGTGGAACCTGAGAGGATATACGATCTGTAGA	2340
Dp	2281	TCTCTGGAATAATTAGAGCGAGAAAGTGGGTGGAACCTGAGAGGATATACGATCTGTAGA	2340
Oy	2341	CAGAGAACTAAGATGACAGATGAAACAGAACAGGGAAGGTATGATTTTCTTGAANAAT	2400
Dp	2341	CAGAGAACTAAGATGACAGATGAAACAGAACAGGGAAGGTATGATTTTCTTGAANAAT	2400
Oy	2401	CACAAACTTGAACATTTTGTATTAACCTACTGATTCACATTTTGTATATATTTGTCCAA	2460
Dp	2401	CACAAACTTGAACATTTTGTATTAACCTACTGATTCACATTTTGTATATATTTGTCCAA	2460
Oy	2461	AAAACTGTGGTGGTTTGAAGATGATGACACACGACAGTTTTCATGATATTCATCAAA	2520
Dp	2461	AAAACTGTGGTGGTTTGAAGATGATGACACACGACAGTTTTCATGATATTCATCAAA	2520
Oy	2521	CAGATCCATGAAGAAAGAGACGCAAAAGAGGGAATTCAGAGATGTTCCACAGAGGAA	2580
Dp	2521	CAGATCCATGAAGAAAGAGACGCAAAAGAGGGAATTCAGAGATGTTCCACAGAGGAA	2580
Oy	2581	CGTGCCAAGTGTGTTGGCCAGACAGCAGACAACTTAATCCCTCTAGCAATGACGATTC	2640
Dp	2581	CGTGCCAAGTGTGTTGGCCAGACAGCAGACAACTTAATCCCTCTAGCAATGACGATTC	2640
Oy	2641	CGAAAAGAGTATATGTAATACTAACTAACTATCCCTGCGCTTTTGTTTTCAAACTTA	2700
Dp	2641	CGAAAAGAGTATATGTAATACTAACTAACTATCCCTGCGCTTTTGTTTTCAAACTTA	2700
Oy	2701	AGAGTAACTGAATTAATTCGGGTTTGAATCTTTTGGCAGAGCTGAGAAAGTGTCAAGCTTC	2760
Dp	2701	AGAGTAACTGAATTAATTCGGGTTTGAATCTTTTGGCAGAGCTGAGAAAGTGTCAAGCTTC	2760
Oy	2761	ATCGAGTTTCAAGAGAAAGAGATGAGAGAGTTTGTGAAGAGAGGAGAGTCTGATAAA	2820
Dp	2761	ATCGAGTTTCAAGAGAAAGAGATGAGAGAGTTTGTGAAGAGAGGAGAGTCTGATAAA	2820
Oy	2821	GATCAGAGAGAGAGATGGAAGACATGAAGAAAGGACATCACAGAGAGATATTTGATCTG	2880
Dp	2821	GATCAGAGAGAGAGATGGAAGACATGAAGAAAGGACATCACAGAGAGATATTTGATCTG	2880
Oy	2881	GAGAAAGAAATTTGAGTGAAGCTTTGGAAACAGTCAAGTCAAGCAATGGCTTCAAAATGA	2940
Dp	2881	GAGAAAGAAATTTGAGTGAAGCTTTGGAAACAGTCAAGTCAAGCAATGGCTTCAAAATGA	2940
Oy	2941	GATGATTTGAGCAAAAAGTCTGTATACACAAGACAAGACTAATTTTCTTTGTTTGGCTTTG	3000
Dp	2941	GATGATTTGAGCAAAAAGTCTGTATACACAAGACAAGACTAATTTTCTTTGTTTGGCTTTG	3000
Oy	3001	GTAATGTCGAAAATGAGAGATCTGAGAGACTTCATTTAATATCTAGCAAAATCTAAGAA	3060

	3060	3120	3180	3240
Db	GTATGTCGAAAGATGAGAGATCTGAGAGACTCCATTTAAATACTAGACCAATCTAAGGA			
Qy	GATATATGATTTATTAATCTCCCAATTTTACTAGACGATCTAAGGAGCAATTAAGTCTT			
Db	GATTATGATTTATTTATCTCCCAATTTTATAGAGCGGATCTAAGGAAAGCAATTAGTCTTT			
Qy	GTGACTTAAACCAAGTTCCTTAGTATTTTGTTTTTTGGTAAAAATTTCATATGAAG			
Db	GTGCTCTAAACCAAGTTCCTTAGTATTTTGTTTTTTGGTAAAAATTTCATATGAAG			
Qy	TTAGACATTTTACCAACGCTCAGAGTAATCAACAGATGGCAAAATCATAGTCTTTT			
Db	TTAGACATTTATCCAAACGCTCAGAGTAAATCAACAGATGGCAAAATCATAGTCTTTT			
Qy	AGAAATTTTATATCTACCAAAATATATGGGTCAAAAT			
Db	AGAAATTTTATATCTACCAAAATATATGGGTCAAAAT			

RESULT 2	
AB025633	AB025633 81365 bp DNA linear PLN 27-DEC-2000
LOCUS	
DEFINITION	Arabidopsis thaliana genomic DNA, chromosome 5, F1 clone: M0M1.
ACCESSION	AB025633 BA000015
VERSION	AB025633.2 GI:10178221
KEYWORDS	
SOURCE	
ORGANISM	Arabidopsis thaliana (thale cress)
REFERENCE	Arabidopsis thaliana Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Rosidiales; Magnoliophyta; eudicotyledons; core eudicots; rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsie. 1 (stee)
AUTHORS	Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H. and Tabata, S.
TITLE	Structural analysis of Arabidopsis thaliana chromosome 5. X. Sequence features of the regions of 3,076,755 bp covered by sixty P1 and TAC clones
JOURNAL	DNA Res. 7 (1), 31-63 (2000)
MEDLINE	20181125
PIRMEED	10718197
REFERENCE	2 (bases 1 to 81365)
AUTHORS	Nakamura, Y.
TITLE	Direct Submission
JOURNAL	Submitted (02-APR-1999) Yasukazu Nakamura, Kazusa DNA Research Institute, Department of Plant Gene Research; 1532-3, Yana, Kisarazu, Chiba 292-0812, Japan (E-mail: ynakamu@kazusa.or.jp, Tel:81-438-52-3935; Fax:81-438-52-3934)
COMMENT	On Sep 15, 2000 this sequence version replaced gi:4589439. Address for correspondence: kase@kazusa.or.jp http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=M0M1 For the latest information on annotation of this clone, please see http://www.kazusa.or.jp/kaos/cgi-bin/aggd_graph.cgi?c=M0M1 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Graal (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/Graal-1.3/), GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html), NetGene2 (S.-M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlin1.zool.laastate.edu/cgi-bin/rp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is K19M13 and the 3' clone is M0011. Location/Qualifiers 1..81365 /organism="Arabidopsis thaliana" /mol_type="genomic DNA"
FEATURES	
SOURCE	

```

/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone_lib="Mitsui P1"
join(3184..3474..3595,3681..3723,3777..3898,
4042..4198,4445..4515,4601..4714,4888..4988,5300..5420)
/notes="gb|AAF01580.1"
gene_id:MOM1.24
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97251.1"
/db_xref="GI:8809710"
/translation="MHWQGTSDVDSPPAKKVDTPPKHVOFLFETSKTVVDNAQNS
YSAFDPSSNSPLLSVFERSSSPKSGDDPLPGIEDLOISEEYPGHELOACG
YSINGTSCNFMVCHLEDGVSNIYIDESSRETSNAGAKOPNYLTADVDLYIAEV
OPLDNRKGRKRFKGRYDVLVDVTLFSDCSTSVFPMWFAIFICFPIISGG
ACDPMQNIETKLTHTGASYSKSLAVGVIVMEATLISIRREGSJKICSDLTIAE
KRSASTVTIPGQPAELVIIGSDSEHSLRADNSPDLIGSRDEIVLTLRFIKVM
FFKALNSFYHFRCSCHKSRNL"
join(6317..6428,6508..6629,6714..6756,6841..6932,
7121..7169,7396..7466,7550..7663,7785..7882,8112..8181)
/notes="gb|AAF01580.1"
gene_id:MOM1.23
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97250.1"
/db_xref="GI:8809709"
/translation="WQGTSVSPALNQADETTPISKTVNDTKNIFFPDGDDAPL
PALENTQISEEYPGHELOACGYSINGTSCNFMVCHLEDGVSNIYIDESSRETSN
ADVGLCLAEVQPLDNRKGRKRYFANPNKRIKCHPEMOSNIDKTLHTGASIKV
SLIGFHWIWEATLISIRREGYTIKCNVDLITFEKFSASTAVKIPFEKPAELVTIGSD
GSEHCARVNDENPDIRLCKEKKVRLRYKYVVFEOKL"
complement(join(6642..8705,8794..8863,8978..9387,
9584..10347))
/notes="gene_id:MOM1.22
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97249.1"
/db_xref="GI:8809708"
/translation="WSMWKGSQGTWLPDLKQKQKGLSEVGEPPPPSTVSNASF
GVRGLRNHSPSEKSFSSVLPSPKALPTEKDCNOERGCGCCRRKPTLSLPNS
HDLATLKXENSWADNLTLDVLTSTEDPEMALATLKQVSGKDEEPTKIRBY
SSDNRSEYTFEKTIVSVMAARSTEDKGYDLKSDGSSFLVNASNEKFPDI
SELDIIORLOSIPIEPEWEDDLYLSHRKDALVKRSASNSHSAQNAQORLDHSA
KQHSKAPEDMLAAEKLENAEAKKIGITNKDNDIMKLDLGLHATAVAVQDLQERLQ
LEGHTVRSVSPNRGRKNAALRSASQEPGRLDDEGMCORTSRELNSLOVITG
IGKHSRQASLPLAVKTFEDNRKRFEDTRPGVTVRPPKFS"
complement(12417..13424)
/notes="contains similarity to unknown protein"
gene_id:MOM1.21"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97248.1"
/db_xref="GI:8809707"
/translation="MATDSQPNQKLTLPKTRIAITVISTNTDNAQRPDGTINRFLR
LDFRAPNPKPKNIVSTSDFVQSDRLWFLYTPHVSQKIPVAVFPGGAFIS
PAAYPDVNCRRFARKLPAYVIVSNVRLAEBHRYPQVDGDLAKYIEHSGSILPA
NADLSRCPAGDSAGNTAHVAIRICREPRSPFAVKLIGLISIPFGGEETEAR
KQLVGAFVISPRTDMCMZMGINDHEAVVNGGNVADISGLDYPETWVAVAGDPL
KQWQSYEWMLKCGKATLIEYPNNFAFYFPLPBAQGLIMRIKQFVDBRVASLS
A"
join(14922..14981,15072..15167,15563..15666,15755..15891,
16096..16178)
/notes="gene_id:MOM1.20"
/codon_start=1
/evidence=not experimental
/protein_id="508_ribosomal_protein_L24"
/notes="contains similarity to unknown protein"
/strain="Columbia"
/db_xref="taxon:3702"
/chromosome="5"
/clone_lib="Mitsui P1"
join(3184..3474..3595,3681..3723,3777..3898,
4042..4198,4445..4515,4601..4714,4888..4988,5300..5420)
/notes="gb|AAF01580.1"
gene_id:MOM1.24
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97251.1"
/db_xref="GI:8809710"
/translation="MHWQGTSDVDSPPAKKVDTPPKHVOFLFETSKTVVDNAQNS
YSAFDPSSNSPLLSVFERSSSPKSGDDPLPGIEDLOISEEYPGHELOACG
YSINGTSCNFMVCHLEDGVSNIYIDESSRETSNAGAKOPNYLTADVDLYIAEV
OPLDNRKGRKRFKGRYDVLVDVTLFSDCSTSVFPMWFAIFICFPIISGG
ACDPMQNIETKLTHTGASYSKSLAVGVIVMEATLISIRREGSJKICSDLTIAE
KRSASTVTIPGQPAELVIIGSDSEHSLRADNSPDLIGSRDEIVLTLRFIKVM
FFKALNSFYHFRCSCHKSRNL"
join(6317..6428,6508..6629,6714..6756,6841..6932,
7121..7169,7396..7466,7550..7663,7785..7882,8112..8181)
/notes="gb|AAF01580.1"
gene_id:MOM1.23
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97250.1"
/db_xref="GI:8809709"
/translation="WQGTSVSPALNQADETTPISKTVNDTKNIFFPDGDDAPL
PALENTQISEEYPGHELOACGYSINGTSCNFMVCHLEDGVSNIYIDESSRETSN
ADVGLCLAEVQPLDNRKGRKRYFANPNKRIKCHPEMOSNIDKTLHTGASIKV
SLIGFHWIWEATLISIRREGYTIKCNVDLITFEKFSASTAVKIPFEKPAELVTIGSD
GSEHCARVNDENPDIRLCKEKKVRLRYKYVVFEOKL"
complement(join(6642..8705,8794..8863,8978..9387,
9584..10347))
/notes="gene_id:MOM1.22
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97249.1"
/db_xref="GI:8809708"
/translation="WSMWKGSQGTWLPDLKQKQKGLSEVGEPPPPSTVSNASF
GVRGLRNHSPSEKSFSSVLPSPKALPTEKDCNOERGCGCCRRKPTLSLPNS
HDLATLKXENSWADNLTLDVLTSTEDPEMALATLKQVSGKDEEPTKIRBY
SSDNRSEYTFEKTIVSVMAARSTEDKGYDLKSDGSSFLVNASNEKFPDI
SELDIIORLOSIPIEPEWEDDLYLSHRKDALVKRSASNSHSAQNAQORLDHSA
KQHSKAPEDMLAAEKLENAEAKKIGITNKDNDIMKLDLGLHATAVAVQDLQERLQ
LEGHTVRSVSPNRGRKNAALRSASQEPGRLDDEGMCORTSRELNSLOVITG
IGKHSRQASLPLAVKTFEDNRKRFEDTRPGVTVRPPKFS"
complement(12417..13424)
/notes="contains similarity to unknown protein"
gene_id:MOM1.21"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97248.1"
/db_xref="GI:8809707"
/translation="MATDSQPNQKLTLPKTRIAITVISTNTDNAQRPDGTINRFLR
LDFRAPNPKPKNIVSTSDFVQSDRLWFLYTPHVSQKIPVAVFPGGAFIS
PAAYPDVNCRRFARKLPAYVIVSNVRLAEBHRYPQVDGDLAKYIEHSGSILPA
NADLSRCPAGDSAGNTAHVAIRICREPRSPFAVKLIGLISIPFGGEETEAR
KQLVGAFVISPRTDMCMZMGINDHEAVVNGGNVADISGLDYPETWVAVAGDPL
KQWQSYEWMLKCGKATLIEYPNNFAFYFPLPBAQGLIMRIKQFVDBRVASLS
A"
join(14922..14981,15072..15167,15563..15666,15755..15891,
16096..16178)
/notes="gene_id:MOM1.20"
/codon_start=1
/evidence=not experimental
/protein_id="508_ribosomal_protein_L24"
/notes="contains similarity to unknown protein"

```

CDS
complement(join(29439..29576,29661..29885,29985..30268,
30341..30493,30579..30722,30956..31484))
/note="gene_id:WOM1.15"

Query Match 100.0%; Score 3275; DB 8; Length 81365;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3275; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAAACAAACAAACAAATTTAGCAAGTCATGTTGTAGCAATAATTAATAGTGGAAACA 60
DB 21977 GACAAACAAACAAACAAATTTAGCAAGTCATGTTGTAGCAATAATTAATAGTGGAAACA 22036
QY 61 TTAAGTTAAGCGAAACAAAGGAAAAAAGGTACAAAAATGAAAAACAACAACTGAAT 120
DB 22037 TTAAGTTAAGCGAAACAAAGGAAAAAAGGTACAAAAATGAAAAACAACAACTGAAT 22096
QY 121 GAAATTTGAGTCAGAAATCGAAAAACGAGGCCGTTTAAATAGCTTCTC 180
DB 22097 GAAATTTGAGTCAGAAATCGAAAAACGAGGCCGTTTAAATAGCTTCTC 22156
QY 181 ATTGTCCTCTTCTGCGAGTTTATTTCTCTCCGAGTCTGACTGCTCTC 240
DB 22157 ATTGTCCTCTTCTGCGAGTTTATTTCTCTCCGAGTCTGACTGCTCTC 22216
QY 241 TCTCCGCGCTTAAACCTTACGTTCTCCGCTTACTCTGAAGTTTCTGCTTAGAG 300
DB 22217 TCTCCGCGCTTAAACCTTACGTTCTCCGCTTACTCTGAAGTTTCTGCTTAGAG 22276
QY 301 CCTCCGATGCTCACCGCATGCAATCTGCTCGAATTTCTTTTCTTCTGCTGAAAA 360
DB 22277 CCTCCGATGCTCACCGCATGCAATCTGCTCGAATTTCTTTTCTTCTGCTGAAAA 22336
QY 361 ATTGCGCTAATGTTCTCGAATTTGAAAGTTTGTGCTAAGGTTAATTTTCCCTAAT 420
DB 22337 ATTGCGCTAATGTTCTCGAATTTGAAAGTTTGTGCTAAGGTTAATTTTCCCTAAT 22396
QY 421 TTTTATAGTTCTTAGGTAACGATACCGCTTACCTTATGTTTGTTCATTTGTGAGCT 480
DB 22397 TTTTATAGTTCTTAGGTAACGATACCGCTTACCTTATGTTTGTTCATTTGTGAGCT 22456
QY 481 TTCACGCTTATGCTGCTGATCGAGATATTGACTGTGAAAAATCCTTCGTTTGTGTT 540
DB 22457 TTCACGCTTATGCTGCTGATCGAGATATTGACTGTGAAAAATCCTTCGTTTGTGTT 22516
QY 541 TTGTTTCAATTAATGATGATGATCTACCTTTTGTGCTTGAATGTTTGTGAGCCT 600
DB 22517 TTGTTTCAATTAATGATGATGATCTACCTTTTGTGCTTGAATGTTTGTGAGCCT 22576
QY 601 ATGCGTTGTTGCTGTTTAACTTCAACGTTCAATGTTGATTTTGAATTTGGTAGTG 660
DB 22577 ATGCGTTGTTGCTGTTTAACTTCAACGTTCAATGTTGATTTTGAATTTGGTAGTG 22636
QY 661 ACTGTGGGTTTCTTGTGCTGCTAATAGTTGTAAAAATGAGTTTCTAGGGCTGTCAATGT 720
DB 22637 ACTGTGGGTTTCTTGTGCTGCTAATAGTTGTAAAAATGAGTTTCTAGGGCTGTCAATGT 22696
QY 721 CTAAAGAAAAAAGACCTTTCAGGGGTGTTAAGGCTTGAACAGCTTGTTCAGAGTT 780
DB 22697 CTAAAGAAAAAAGACCTTTCAGGGGTGTTAAGGCTTGAACAGCTTGTTCAGAGTT 22756
QY 781 TGGCAGGAGACAGACTGCTTCTTCAACAGATGATGAGAGAGTGGAGGTCAATTCCA 840
DB 22757 TGGCAGGAGACAGACTGCTTCTTCAACAGATGATGAGAGAGTGGAGGTCAATTCCA 22816
QY 841 AGAAGAACAAAGAACAAACAGAAACACTTCTGAAAAAACTTGGGTTTCTCAGAAATTGCA 900
DB 22817 AGAAGAACAAAGAACAAACAGAAACACTTCTGAAAAAACTTGGGTTTCTCAGAAATTGCA 22876
QY 901 ATCTCTCTAGAGTTGGGGTGTCTACAGACAGAGAGAGATAGACATCTGGAGAG 960
DB 22877 ATCTCTCTAGAGTTGGGGTGTCTACAGACAGAGAGAGATAGACATCTGGAGAG 22936
QY 961 GAAACAAATGATTCGGGAGAGTTAACGCAATGCTGGGGCAATTCAAGCTTAACATATCTG 1020

DB 22937 GAAACAAATGATTCGGGAGAGTTAACGCAATGCTGGGGCAATTCAAGCTTAACATATCTG 22996
QY 1021 GTCCGGGAGCAGCGTTGAGCAGAAAAGTATGATTAACAATCTTTGTGAGCAACCCCACTGTAT 1080
DB 22997 GTCCGGGAGCAGCGTTGAGCAGAAAAGTATGATTAACAATCTTTGTGAGCAACCCCACTGTAT 23056
QY 1081 CTCGCCCTCCTTTGAAAGAGATGGAATTTGGCAGGCAAGAGAGGTCTGCTCAGCA 1140
DB 23057 CTCGCCCTCCTTTGAAAGAGATGGAATTTGGCAGGCAAGAGAGGTCTGCTCAGCA 23116
QY 1141 CAGCTGTGAGAGGTTTCTCTGACGTGAGAGATGATGTGATTAATGCTTCTGAGAGAGA 1200
DB 23117 CAGCTGTGAGAGGTTTCTCTGACGTGAGAGATGATGTGATTAATGCTTCTGAGAGAGA 23176
QY 1201 ATGATTCGATGCTTTGATGATATCTGATGACGACTTCGCAAGTATGATGACTCGG 1260
DB 23177 ATGATTCGATGCTTTGATGATATCTGATGACGACTTCGCAAGTATGATGACTCGG 23236
QY 1261 ATGTGATGAAAAGGCCATGATCAAGAAACAGAAATTAAGGTTCAAAAAGTCTTGG 1320
DB 23237 ATGTGATGAAAAGGCCATGATCAAGAAACAGAAATTAAGGTTCAAAAAGTCTTGG 23296
QY 1321 GAGCTTGTATAGCTTGTGATCGAGCAGATTAATGAACACAGAGCAGTGGCATTTGTC 1380
DB 23297 GAGCTTGTATAGCTTGTGATCGAGCAGATTAATGAACACAGAGCAGTGGCATTTGTC 23356
QY 1381 CAGCTTGTGAAAACGACCTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 1440
DB 23357 CAGCTTGTGAAAACGACCTGTGCTGATGATGATGATGATGATGATGATGATGATGATG 23416
QY 1441 ATGCGAGGACAAAGAGGCTAGGCGAGTAAGCTCATGAGAAATGCGTGAAGTTTAG 1500
DB 23417 ATGCGAGGACAAAGAGGCTAGGCGAGTAAGCTCATGAGAAATGCGTGAAGTTTAG 23476
QY 1501 AAAAGATCTACAGATGAGAGCGCATCTGTCAATCTTGTGTGAGATTTAGGCGAGT 1560
DB 23477 AAAAGATCTACAGATGAGAGCGCATCTGTCAATCTTGTGTGAGATTTAGGCGAGT 23536
QY 1561 GGAAGGTTTGGGTGAGATGAAAAAGATTAAGAAATTTGTGCGCTCCAAATGCTCATCA 1620
DB 23537 GGAAGGTTTGGGTGAGATGAAAAAGATTAAGAAATTTGTGCGCTCCAAATGCTCATCA 23596
QY 1621 TCATGAATCTAGACTGCTAAGAGCAGATAAGATTAAGTGAATTTCTCTGCTTTTAC 1680
DB 23597 TCATGAATCTAGACTGCTAAGAGCAGATAAGATTAAGTGAATTTCTCTGCTTTTAC 23656
QY 1681 TTCTTAATTTTCTCTGCTCAATCTAGATCTTGAATGTTTACATGATGATGATGATGATG 1740
DB 23657 TTCTTAATTTTCTCTGCTCAATCTAGATCTTGAATGTTTACATGATGATGATGATGATG 23716
QY 1741 ATGGGCAACCAAGAGCTGCTGAAATTAAGTCAACAGATTAAGAGCTTTAGACACGCCAT 1800
DB 23717 ATGGGCAACCAAGAGCTGCTGAAATTAAGTCAACAGATTAAGAGCTTTAGACACGCCAT 23776
QY 1801 TCCATATGCTCCACAGGCGCATGCTGAGATGATGATGATGATGATGATGATGATGATGATG 1860
DB 23777 TCCATATGCTCCACAGGCGCATGCTGAGATGATGATGATGATGATGATGATGATGATGATG 23836
QY 1861 GGCATATTTGAGAGCGCAACGCTCAACCGGAGTTAGCTGATGAGGTTAATAGAAAT 1920
DB 23837 GGCATATTTGAGAGCGCAACGCTCAACCGGAGTTAGCTGATGAGGTTAATAGAAAT 23896
QY 1921 GCCTGGGGTCAAGACGCGATATGTTTCTGAGAGTTTCCCAACTGATAGGCTTCTT 1980
DB 23897 GCCTGGGGTCAAGACGCGATATGTTTCTGAGAGTTTCCCAACTGATAGGCTTCTT 23956
QY 1981 GGAAGAACCAAGATCTGACATTAATCAACATCTGCAAGTCTCTCCCAAGAGA 2040
DB 23957 GGAAGAACCAAGATCTGACATTAATCAACATCTGCAAGTCTCTCCCAAGAGA 24016
QY 2041 AATTTGATTAATGCTTTAGTTTGTCAATGGAATTTAAAGTTTGTGCTCCGTTTAA 2100

```

Db 24017 AATTGATATATGCTTTAGTTTGTCAATGGAATTTAAAGTTTGTGTCGTTGTA 24076
QY 2101 TGCATCTGTATGATATATATATATATATATATATATATATATATATATATAT 2160
Db 24077 TGCATCTGTATGATATATATATATATATATATATATATATATATATATATAT 24136
QY 2161 AATCATACCAAGAGATGTTGTTGAAGAGATGAGGAGATCTCTGAGAGACATCAGAG 2220
Db 24137 AATCATACCAAGAGATGTTGTTGAAGAGATGAGGAGATCTCTGAGAGACATCAGAG 24196
QY 2221 CTGAACCTATTAAGAACAGCTCTCAAAAACAGAACAGGAGGAGGAGGAGGAG 2280
Db 24197 CTGAACCTATTAAGAACAGGCTCTCAAAAACAGAACAGGAGGAGGAGGAGGAG 24256
QY 2281 TCTCTGGAATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2340
Db 24257 TCTCTGGAATTAATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 24316
QY 2341 CAGAGAACTAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2400
Db 24317 CAGAGAACTAAGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 24376
QY 2401 CACAAACCTGACATTTTGTATTAATCTAATGATTCACATTTTGTATTAATG 2460
Db 24377 CACAAACCTGACATTTTGTATTAATCTAATGATTCACATTTTGTATTAATG 24436
QY 2461 AAAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2520
Db 24437 AAAACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 24496
QY 2521 CAGATCATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2580
Db 24497 CAGATCATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24556
QY 2581 CGTGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2640
Db 24557 CGTGCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 24616
QY 2641 CAAAAGAGGATATATGATATATATATATATATATATATATATATATATATATAT 2700
Db 24617 CAAAAGAGGATATATGATATATATATATATATATATATATATATATATATATAT 24676
QY 2701 AAGATTAATGAAATTTATTCGGTTTGTATTCCTTTCGAGAGCTGAGAGAGTCAAGCTTC 2760
Db 24677 AAGATTAATGAAATTTATTCGGTTTGTATTCCTTTCGAGAGCTGAGAGAGTCAAGCTTC 24736
QY 2761 ATCGAGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2820
Db 24737 ATCGAGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24796
QY 2821 GATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2880
Db 24797 GATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24856
QY 2881 GAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2940
Db 24857 GAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24916
QY 2941 GATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3000
Db 24917 GATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24976
QY 3001 GATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3060
Db 24977 GATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25036
QY 3061 GATTAAGATTAATATCTCAATTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 3120
Db 25037 GATTAAGATTAATATCTCAATTTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 25096
QY 3121 GTGACTAAACCAAGTTCTCTAGATATATATATATATATATATATATATATATATATAT 3180
Db 25097 GTGACTAAACCAAGTTCTCTAGATATATATATATATATATATATATATATATATATAT 25156

```

```

QY 3181 TTAGACATATTAACCAAGGTCAGAGTGAATCAAGAAATGGAATCAAAATCATGTTTTT 3240
Db 25157 TTAGACATATTAACCAAGGTCAGAGTGAATCAAGAAATGGAATCAAAATCATGTTTTT 25216
QY 3241 AGAATTTTATATCTACAAATATATATGAGTACAAAT 3275
Db 25217 AGAATTTTATATCTACAAATATATATGAGTACAAAT 25251

RESULT 3
AF239719 2254 bp DNA linear PLN 02-JUN-2000
LOCUS Arabidopsis thaliana SGS3 gene, complete cds.
DEFINITION Arabidopsis thaliana SGS3 gene, complete cds.
ACCESSION AF239719
VERSION AF239719.1 GI:8164029
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 2254)
Mourrain, P., Becil, C., Elmayan, T., Feuerbach, F., Godon, C.,
Morel, J.B., Jouette, D., Lacombe, A.M., Nikic, S., Picault, N.,
Remoue, K., Sanial, M., Vo, T.A. and Vaucheret, H.
Arabidopsis SGS2 and SGS3 genes are required for
posttranscriptional gene silencing and natural virus resistance
Cell 101 (5), 533-542 (2000)
10850495
MEDLINE
JOURNAL
PUBMED
REFERENCE
2 (bases 1 to 2254)
Becil, C., Mourrain, P., Vaucheret, H. and Elmayan, T.
Direct Submission
Submitted (28-FEB-2000) Biologie Cellulaire, INRA, Route de
Saint-Cyr, Versailles 78026, France
FEATURES
source
1..2254.
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/chromosome="v"
join(<1..963,1037..1328,1440..1684,1787..1953,2044..>2254)
/product="SGS3"
join(1..963,1037..1328,1440..1684,1787..1953,2044..2254)
/codon_start=1
/product="SGS3"
/protein_id="AA073960.1"
/db_xref="GI:8164030"
/translation="MSSRAGPMKSRKRVGGYRPREVQLVQGLAETRLASSODDGGEM
EVIISKKNKRGNTSGKTVVSONSIPPAKMGQOQGRSSNYSGRNNVSGKNGNGRG
IQAITSGRRLSRKYDNNFVAPEVSPPLEGGNMGWARGSAQHTAVQEPVDED
VDNASEENDSDALDDSDDLASDVIDVQSKSGSRKQKWFKKFGLDLSLEIQ
INEPQRIQHWPCACQNGPGLIDMYNLHPLAATITGKARVYLHRLAFLAEKDLQMRG
ASVIEQGEIYQWKGLEDEKQYELVPEPILINMTLRDNDNDKMLQNGOELLEYF
DKVELRLRHSHYGPQGHGMSVIMESSATYLELRHLRELAEGLRIAMGQKRM
FSGGRQLYGLPLATQDDIDIPNHSQGTRLFELKSYQEMVYELROISDNOOLY
FNKLSKQNKIAFLVESLEITSEKLRATADNRIVRQRTKQHEQNEEMDAHDFE
MDSIQIHERPDAAKEENPEMLQOERARVVOQOQNNIPSSNDRCRKAEEVSSIEF
QEKEMEEVEREREMLIKQEKEMEDMKRRHHEIFIDLEKEFDEALEOLMYKHGLHNE
D"

ORIGIN
Query Match 68.8%; Score 2254; DB 8; Length 2254;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2254; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 696 ATGAGTTCTAGGGCTGTCGTCATGCTAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 755
Db 1 ATGAGTTCTAGGGCTGTCGTCATGCTAAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 60
QY 756 GAGTTGAACAGTTGTTCAAGGTTTGCAGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 815

```


|||||
Db 61 GAGGTGAACAGTGGTTCAAGTTGGAGGAGCAGACTGGCTCTTCAACAATGAT .120
|||
Qy 816 GGAGGAGTGGGAGGTCATTTCCAGAAAGAACAAACACAGAGAAACCTTCTGA 875
|||
Db 121 GGAGGAGTGGGAGGTCATTTCCAGAAAGAACAAACACAGAGAAACCTTCTGA 180
|||
Qy 876 AAAACTTGGGTTCTCAGAAATTCGATCTCTAGAGCTTGGGGGTGCAGCAGAGG 935
|||
Db 181 AAACTTGGGTTCTCAGAAATTCGATCTCTAGAGCTTGGGGGTGCAGCAGAGG 240
|||
Qy 936 AGAGTAGCAACGTATCTGGAGAGAGAAACAATGATCCGGAGAGGTAAACCGCAATGGT 995
|||
Db 241 AGAGTAGCAACGTATCTGGAGAGAGAAACAATGATCCGGAGAGGTAAACCGCAATGGT 300
|||
Qy 996 CCGGGCATTCAGACTAACATATTCGGTCCGGGACGAGCCTTGAGCAGAAAGTATGATTAAC 1055
|||
Db 301 CCGGGCATTCAGACTAACATATTCGGTCCGGGACGAGCCTTGAGCAGAAAGTATGATTAAC 360
|||
Qy 1056 AACTTGTGGCAACCCACCTGATCTCGCCCTCTTGGAGAGAGATGGAATTGGCAG 1115
|||
Db 361 AACTTGTGGCAACCCACCTGATCTCGCCCTCTTGGAGAGAGATGGAATTGGCAG 420
|||
Qy 1116 GCAAGAGAGGTTCTGCTCAGCAACAGCTGTGACAGAGTTTCTGACGTGAGATGAT 1175
|||
Db 421 GCAAGAGAGGTTCTGCTCAGCAACAGCTGTGACAGAGTTTCTGACGTGAGATGAT 480
|||
Qy 1176 GTGATTAATGCTTCTGAGAGAGATGATCCGATGCTTTGATGATGATGACGAC 1235
|||
Db 481 GTGATTAATGCTTCTGAGAGAGATGATCCGATGCTTTGATGATGATGACGAC 540
|||
Qy 1236 CTTCGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1295
|||
Db 541 CTTCGAAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
|||
Qy 1296 AATTAAGTGTCAAAAAGTCTTTGGCAGCTTGATGATGATGATGATGATGATGATGATGAT 1355
|||
Db 601 AATTAAGTGTCAAAAAGTCTTTGGCAGCTTGATGATGATGATGATGATGATGATGATGAT 660
|||
Qy 1356 GAACCAACAAGCAGTGGCATTGTCAGCTGTGACAAAGCAGCTGTGTCATGATGATGAT 1415
|||
Db 661 GAACCAACAAGCAGTGGCATTGTCAGCTGTGACAAAGCAGCTGTGTCATGATGATGATGAT 720
|||
Qy 1416 TATTAACCTGACCCCTTAAGCTCATGAGGACAAAGAGCTAGAGGCAATTCATGATGAT 1475
|||
Db 721 TATTAACCTGACCCCTTAAGCTCATGAGGACAAAGAGCTAGAGGCAATTCATGATGATGAT 780
|||
Qy 1476 CATAGAATTTGGCTGAAGTTTGAAGAAAGATCTACAGATGAGAGCGCATCTGATGAT 1535
|||
Db 781 CATAGAATTTGGCTGAAGTTTGAAGAAAGATCTACAGATGAGAGCGCATCTGATGATGAT 840
|||
Qy 1536 CCTTGTGGTGAAGTTTATGGGAGTGGAGAGGTTTGGTGAAGATGATGATGATGATGAT 1595
|||
Db 841 CCTTGTGGTGAAGTTTATGGGAGTGGAGAGGTTTGGTGAAGATGATGATGATGATGATGAT 900
|||
Qy 1596 ATTGTCGGCTCCATGAGTCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
|||
Db 901 ATTGTCGGCTCCATGAGTCATCATGATGATGATGATGATGATGATGATGATGATGATGAT 960
|||
Qy 1656 AAGGTGGAATTTCTGCTCTTACTTCTTAAATTTTCTTGGCAATCTACTGATCTTAA 1715
|||
Db 961 AAGGTGGAATTTCTGCTCTTACTTCTTAAATTTTCTTGGCAATCTACTGATCTTAA 1020
|||
Qy 1716 GAATGTTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1775
|||
Db 1021 GAATGTTAATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
|||
Qy 1776 GTATGAGGCTTTAGAGCAGCATCTATGATGTCACAGGGCAGCATGCGGAGTATGATGAT 1835
|||
Db 1081 GTATGAGGCTTTAGAGCAGCATCTATGATGTCACAGGGCAGCATGCGGAGTATGATGATGAT 1140
|||
Qy 1836 TGTGATGTTGAGAGAGTGCATGCGCTATTTGAGGCGGAAACGCTTCACCGGAGATT 1895
|||

|||||
Db 1141 TGTGATGTTGAGAGCAGTGCACCTGCTATTTGAGAGCCGCTCCACCGGAGATT 1200
|||
Qy 1896 AGCTGAGATGGGTTAGATAGAAATTTGCCCTGGGGTCAGAGCGCAGATGTTTCTGAGAG 1955
|||
Db 1201 AGCTGAGATGGGTTAGATAGAAATTTGCCCTGGGGTCAGAGCGCAGATGTTTCTGAGAG 1260
|||
Qy 1956 TGTTCGCCAATGTAATGCTTCTTCCAGAGAGAGATCTGACATATTCATCAACA 2015
|||
Db 1261 TGTTCGCCAATGTAATGCTTCTTCCAGAGAGAGATCTGACATATTCATCAACA 1320
|||
Qy 2016 CTCTCAAGGTTCTCTCCCAAGAAATTTGATATATGCTTTAGTTTGTCAATGGAAT 2075
|||
Db 1321 CTCTCAAGGTTCTCTCCCAAGAAATTTGATATATGCTTTAGTTTGTCAATGGAAT 1380
|||
Qy 2076 TTAAGTTTGTGATGCTGCTGATTAATGATCTGATATGATATCTATGATTAATGATGAT 2135
|||
Db 1381 TTAAGTTTGTGATGCTGCTGATTAATGATCTGATATGATATCTATGATTAATGATGAT 1440
|||
Qy 2136 CAAAACAAGGCTGAATTCGATGAAATCATACCAAGAGATGTTGTAAGAGCTGAG 2195
|||
Db 1441 CAAAACAAGGCTGAATTCGATGAAATCATACCAAGAGATGTTGTAAGAGCTGAG 1500
|||
Qy 2196 GCAATCTCTGAGGCAATCAGCAGCTGAATCTTTAAGAACAACTCTCAAAACAGAA 2255
|||
Db 1501 GCAATCTCTGAGGCAATCAGCAGCTGAATCTTTAAGAACAACTCTCAAAACAGAA 1560
|||
Qy 2256 CAGACCGCAGAGGCTTGAAGATCTCTGGAATTTTGAACGAGAGCTGCTGAAGAC 2315
|||
Db 1561 CAGACCGCAGAGGCTTGAAGATCTCTGGAATTTTGAACGAGAGCTGCTGAAGAC 1620
|||
Qy 2316 TGCAGAGATTAATTCGATTCGTGAGACAGAGAACTAAGATGACATGACAGAGAGGGA 2375
|||
Db 1621 TGCAGAGATTAATTCGATTCGTGAGACAGAGAACTAAGATGACATGACAGAGAGGGA 1680
|||
Qy 2376 AAGGATGATTTTCTAGAAATCACAAACTTGAATTTGATTAATCTAATGATGAT 2435
|||
Db 1681 AAGGATGATTTTCTAGAAATCACAAACTTGAATTTGATTAATCTAATGATGATGAT 1740
|||
Qy 2436 CATTTTGAATTAATTTGCCAACAATAAATCTGTGTGTTGAGATGATGATGATGATGATGAT 2495
|||
Db 1741 CATTTTGAATTAATTTGCCAACAATAAATCTGTGTGTTGAGATGATGATGATGATGATGAT 1800
|||
Qy 2496 CAGTTTTTCAATGATTCATCAACAAGATCCATGAAAGAGAGAGAGAGAGAGAGAGAGAG 2555
|||
Db 1801 CAGTTTTTCAATGATTCATCAACAAGATCCATGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 1860
|||
Qy 2556 TTTGAGATGTTGACAGACAGAGAAAGTGCAGAGGTTTGGCCAGACAGACAGAAACAT 2615
|||
Db 1861 TTTGAGATGTTGACAGACAGAGAAAGTGCAGAGGTTTGGCCAGACAGACAGAAACAT 1920
|||
Qy 2616 TAAATCCCTTACGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2675
|||
Db 1921 TAAATCCCTTACGATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1980
|||
Qy 2676 CTGGCGTTTTTCTTCAAACTTAAGATTAATTAATCCGTTTTGATGATGATGATGATGAT 2735
|||
Db 1981 CTGGCGTTTTTCTTCAAACTTAAGATTAATTAATCCGTTTTGATGATGATGATGATGATGAT 2040
|||
Qy 2736 CAGAGCTGAGAGAGTGCAGAGCTTCATGAGTTTCAAGAGAAAGATGAGAGAGTTGAT 2795
|||
Db 2041 CAGAGCTGAGAGAGTGCAGAGCTTCATGAGTTTCAAGAGAAAGATGAGAGAGTTGAT 2100
|||
Qy 2796 GGAAGAGAGAGAGATGCTGATTAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2855
|||
Db 2101 GGAAGAGAGAGAGATGCTGATTAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2160
|||
Qy 2856 GCATCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2915
|||
Db 2161 GCATCAAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2220
|||
Qy 2916 GTACAAGCATGGCTTCACAATGAAGATGATGA 2949
|||
Db 2221 GTACAAGCATGGCTTCACAATGAAGATGATGA 2254
|||

RESULT 4
LOCUS BT002944
DEFINITION Arabidopsis thaliana clone RAF14-93-K05 (R20243) unknown protein
(At5g23570) mRNA, complete cds.
ACCESSION BT002944
VERSION BT002944
KEYWORDS GI:27754622
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Bukeriyota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsids.
1 (bases 1 to 2162)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimu,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.
Arabidopsis Full Length cDNA Clones
Unpublished
2 (bases 1 to 2162)
Yamada,K., Chan,M.M., Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M.,
Onodera,C.S., Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C.,
Yu,G., Yuan,S., Carninci,P., Chen,H., Cheuk,R., Hayashizaki,Y.,
Ishida,J., Jones,T., Kamiya,A., Kawai,J., Kim,C.J., Narusaka,M.,
Nguyen,M., Palm,C.J., Sakurai,T., Satou,M., Seki,M., Shimu,P.,
Southwick,A., Tripp,M.G., Wu,T., Shinozaki,K., Davis,R.W.,
Ecker,J.R. and Theologis,A.
Direct Submission
Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN
Arabidopsis Full-length cDNA"): Seki,M., Narusaka,M., Ishida,J.,
Satou,M., Kamiya,A., Sakurai,T., Carninci,P., Kawai,J.,
Hayashizaki,Y. and Shinozaki,K.
The Salk, Stanford, PGEBC (SSP) Consortium members carried out the
sequencing and annotation of the RAFL cDNAs: Yamada,K., Chan,M.M.,
Chang,C.H., Dale,J.M., Hsuan,V.W., Lee,J.M., Onodera,C.S.,
Quach,H.L., Tang,C.C., Toriumi,M., Wong,C., Wu,H.C., Yu,G.,
Yuan,S., Chen,H., Cheuk,R., Jones,T., Kim,C.J., Nguyen,M.,
Palm,C.J., Shimu,P., Southwick,A., Tripp,M.G., Wu,T., Davis,R.W.,
Ecker,J.R. and Theologis,A.
Yamada,K. (SSP/PGEBC) and Seki,M. (RIKEN GSC) contributed equally to
this work. Shinozaki,K. (RIKEN GSC) and Theologis,A. (SSP/PGEBC)
contributed equally to this work as PIs.
Annotation based on July 2002 version of the Arabidopsis genome
submitted to Genbank.
Location/Qualifiers
1..2162
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAF14-93-K05 (R20243)"
/ecotype="Columbia"
/note="This clone is in a modified pluescript vector2
(lambda PS) as a BamHI/XhoI insert."
1..2162
/gene="At5g23570"
1..91
/gene="At5g23570"
92..1969
/gene="At5g23570"
/codon_start=1

ORIGIN
Query Match 50.4%; Score 1651.2; DB 8; Length 2162;
Best Local Similarity 84.5%; Pred. No. 0;
Matches 2069; Conservative 0; Mismatches 3; Indels 376; Gaps 4;
misc_difference 2146
/evidence=experimental
/product="unknown protein"
/protein_id="AA02257.1"
/db_xref="GI:27754623"
/translation="MSSRAGPMSKKNVQGRPEVEQIVQGLAGTRLASSDDGSEW
/EVISKIKNKPSTGKTWSONSNPRMAGQDQGRGNSVSRGNVSGRGKRG
IOANI SGRGALRYKYDNFVAPPVSRPLEGNNWQARQSAQHTVAQEPDVED
VDNASEENDSDALDSDDDLDSDYDSDVPSLEGNSRKNQKFKFPGSLDLSIEQ
INPQRMHPCACONPGAIIDWVNLHPLAHARTKARVYKLEELAEVLEKQWRG
ASVTPCEITYGKCGEDPKDEIYVPMVITIMNRDLKDDNDKXLMGNOBLEEF
DKYALNARHSYGPQGRGWSVLMFESSATGYI EARIIRLELAEMGLDRIANGCRSM
FSGGVRLVGLATKODLDFNHSQGRPLKLELSYQEMVKEI RQISENQOANM
FKNLSKONKGAIVLESLEIMSEKLRADNRIVRQTKWHEQNRSEMDADHDF
MDSIKOIHERRDKKEENFEMLOQERAKVVGQOQONINPSNDCKRKEVEVSFLF
DEKMEHFVEREMLIKDEKMEKMDKKRHEEIFLEKFPDRLQMLWKGLHNEF
D"
1970..2162
/gene="At5g23570"
misc_difference 2146
/gene="At5g23570"
/note="not present in genomic sequence"
QY 687 GTTGAATAAATGAGTCTTGAAGGCTGTCATGCTTAAGAAAGAAAGCTTCAAGGTGCT 746
DB 83 GTTGAATAAATGAGTCTTGAAGGCTGTCATGCTTAAGAAAGAAAGCTTCAAGGTGCT 142
QY 747 TATAGGCTGAGGTGGAACAGTGTTCAGAGTTTGCGAGGACGACGACTGCTTCTTCA 806
DB 143 TATAGGCTGAGGTGGAACAGTGTTCAGAGTTTGCGAGGACGACGACTGCTTCTTCA 202
QY 807 CAGATGATGAGAGAGATGCGGAGGTCTTTCAGAAAGAAACAGAAACAGAAAC 866
DB 203 CAGATGATGAGAGAGATGCGGAGGTCTTTCAGAAAGAAACAGAAACAGAAAC 262
QY 867 ACTTCGGAATAAATCTGGGTTTCTCAGAAATCGAATCCTCTGAGAGCTTGGGGTGTGAG 926
DB 263 ACTTCGGAATAAATCTGGGTTTCTCAGAAATCGAATCCTCTGAGAGCTTGGGGTGTGAG 322
QY 927 CAGCAAGGAGAGAGTGAACAGTATCTGAGAGAAACAAATGTTCCGGAGAGGTAAAC 986
DB 323 CAGCAAGGAGAGAGTGAACAGTATCTGAGAGAAACAAATGTTCCGGAGAGGTAAAC 382
QY 987 GCGAATGCTCGGGGATTCAGACTAATCTGCTGCGGAGCGAGCGTTGAGCAGAAAG 1046
DB 383 GCGAATGCTCGGGGATTCAGACTAATCTGCTGCGGAGCGAGCGTTGAGCAGAAAG 442
QY 1047 TATGATTAACAATCTTGTGCAACCCCACTGATCTGCGCTCTTGTGAAGAGAGATGG 1106
DB 443 TATGATTAACAATCTTGTGCAACCCCACTGATCTGCGCTCTTGTGAAGAGAGATGG 502
QY 1107 AATTGCGAGGCAAGAGAGAGTCTGCTCAGCAACAAGCTGCGAGAGAGTTCCTGACGG 1166
DB 503 AATTGCGAGGCAAGAGAGAGTCTGCTCAGCAACAAGCTGCGAGAGAGTTCCTGACGG 562
QY 1167 GAGCATGATGAGATTAATGCTCTGAGGAAGAGATGATTCGATGCTTGGATGATTTCT 1226
DB 563 GAGCATGATGAGATTAATGCTCTGAGGAAGAGATGATTCGATGCTTGGATGATTTCT 622
QY 1227 GATGACGACCTTGCAAGTATGATTAATGATCTGAGTGTGATCAAAAGCCATGATTA 1286
DB 623 GATGACGACCTTGCAAGTATGATTAATGATCTGAGTGTGATCAAAAGCCATGATTA 682
QY 1287 CGAAAGCAGAAATAGGTGTTCAAAAGCTTCTGAGGCTTGGATGATGCTTGTGATGAG 1346
DB 683 CGAAAGCAGAAATAGGTGTTCAAAAGCTTCTGAGGCTTGGATGATGCTTGTGATGAG 742
QY 1347 CAGTAAATGAACCAAGAGGACAGTGTGCTGAGTGTTCAGAACGAGACCTGTGTGCC 1406

Db 743 CAGATAAATGACCAAGAGGAGTGGCATTTGTCAGCTTGTGAGAACGAGCCTGTGCC 802
 QY 1407 ATGATTTGGTATTAACCTGCACCCCTCTACTAGCTCATGCGAGACAAAGAGGCTAGGCGA 1466
 Db 803 ATGATTTGGTATTAACCTGCACCCCTCTACTAGCTCATGCGAGACAAAGAGGCTAGGCGA 862
 QY 1467 GTTAACTCTCATAGAGAAATTTGGCTGAAATTTTGAAGAAAGAGCTTACAGATGAGAGCGCA 1526
 Db 863 GTTAACTCTCATAGAGAAATTTGGCTGAAATTTTGAAGAAAGAGCTTACAGATGAGAGCGCA 922
 QY 1527 TCTGTCTATCTTGTGTGTGAGATTTTATGGGAGTGGAGAGGTTTGGGTGAGATGAAAG 1586
 Db 923 TCTGTCTATCTTGTGTGTGAGATTTTATGGGAGTGGAGAGGTTTGGGTGAGATGAAAG 982
 QY 1587 GATTATGAATTTGTGGGCTCCAAATGGTATCATCATGAACTAGACTGATTAAGGAC 1646
 Db 983 GATTATGAATTTGTGGGCTCCAAATGGTATCATCATGAACTAGACTGATTAAGGAC 1042
 QY 1647 GATTATGAATTTGTGGGCTCCAAATGGTATCATCATGAACTAGACTGATTAAGGAC 1706
 Db 1043 GATTATGAATTTGTGGGCTCCAAATGGTATCATCATGAACTAGACTGATTAAGGAC 1052
 QY 1707 CTGATCTTGAATTTTACATTTGATGTGCTGGCATGGGCAACCAAGCTGTGAAATA 1766
 Db 1053 CTGATCTTGAATTTTACATTTGATGTGCTGGCATGGGCAACCAAGCTGTGAAATA 1089
 QY 1767 CTGATCTTGAATTTTACATTTGATGTGCTGGCATGGGCAACCAAGCTGTGAAATA 1826
 Db 1090 CTGATCTTGAATTTTACATTTGATGTGCTGGCATGGGCAACCAAGCTGTGAAATA 1149
 QY 1827 GATGAGTGTCTGATTTTGAAGAGAGTGGCACTGGCTATTTGAGAGCGCAACGCTTCCA 1886
 Db 1150 GATGAGTGTCTGATTTTGAAGAGAGTGGCACTGGCTATTTGAGAGCGCAACGCTTCCA 1209
 QY 1887 CCGGAGTGTGATTTTGAAGAGAGTGGCACTGGCTATTTGAGAGCGCAACGCTTCCA 1946
 Db 1210 CCGGAGTGTGATTTTGAAGAGAGTGGCACTGGCTATTTGAGAGCGCAACGCTTCCA 1269
 QY 1947 TTTCTGAGAGTGTGCGCACTGATGAGGCTTCTTGAAGAGAGAGATGTGACATATT 2006
 Db 1270 TTTCTGAGAGTGTGCGCACTGATGAGGCTTCTTGAAGAGAGAGATGTGACATATT 1329
 QY 2007 CAATCAACACTCTCAAGGTTCTCTCCCAAGAAATTTGATATGATCTTTATTTTGT 2066
 Db 1330 CAATCAACACTCTCAAGGTTCTCTCCCAAGAAATTTGATATGATCTTTATTTTGT 1344
 QY 2067 CATTTGAATTTAAAGTTTGTGTGCTGGCTGATGATCTGTTATGATATCTATGA 2126
 Db 1345 CATTTGAATTTAAAGTTTGTGTGCTGGCTGATGATCTGTTATGATATCTATGA 1344
 QY 2127 TTTATGAGCAAAAGGCTGAAATTTGAGTTGAATCAATCAAGAGAGGTTGTATAA 2186
 Db 1345 TTTATGAGCAAAAGGCTGAAATTTGAGTTGAATCAATCAAGAGAGGTTGTATAA 1398
 QY 2187 GGAGCTGAGGAGATCTTGAGAGCAATGAGCAGTGAATTAATTTAAGAAAGCTCTC 2246
 Db 1399 GGAGCTGAGGAGATCTTGAGAGCAATGAGCAGTGAATTAATTTAAGAAAGCTCTC 1458
 QY 2247 AAAACAGAAACAAGCAGCAGAGTCTTGAAGAAATTTGAGAAATTTATGAGCAGAGCT 2306
 Db 1459 AAAACAGAAACAAGCAGCAGAGTCTTGAAGAAATTTGAGAAATTTATGAGCAGAGCT 1518
 QY 2307 GGGTGAATCTGAGAGATTAATGAGTCTGAGAGCAAGAACTAAGATCAGATGAATAA 2366
 Db 1519 GGGTGAATCTGAGAGATTAATGAGTCTGAGAGCAAGAACTAAGATCAGATGAATAA 1578
 QY 2367 GAAACAGGAGAGTATTTTCTTGAAGAAATCAAAACTGACATTTTGTATTACT 2426
 Db 1579 GAAACAGGAGAGTATTTTCTTGAAGAAATCAAAACTGACATTTTGTATTACT 1589
 QY 2427 ACTGATTCATTTTGTATTAATTTGTCCAAACAAAACCTGTGTGTTGAAGATGA 2486
 Db 1590 ACTGATTCATTTTGTATTAATTTGTCCAAACAAAACCTGTGTGTTGAAGATGA 1596

QY 2487 TGCACACGACAGGTTTTTCATGATTCATCAACAGATCCATGAAAGAGAGACGCAAA 2546
 Db 1597 TGCACACGACAGGTTTTTCATGATTCATCAACAGATCCATGAAAGAGAGACGCAAA 1656
 QY 2547 GGAGAGAAATTTTGAGATTTTGCAGACAGAGAAAGTGCAGAGTTTGTGCGCAGACGA 2606
 Db 1657 GGAGAGAAATTTTGAGATTTTGCAGACAGAGAAAGTGCAGAGTTTGTGCGCAGACGA 1716
 QY 2607 GCAGAAATTTTCCCTTACGAAATGACATTTGCCGAAAGAGATATATGACTAATAAC 2666
 Db 1717 GCAGAAATTTTCCCTTACGAAATGACATTTGCCGAAAGAGATATATGACTAATAAC 1756
 QY 2667 ATATATCTCTGAGGTTTTTGTGTTTCAAACTAAGATGATTAATTCGGTTTTG 2726
 Db 1757 ATATATCTCTGAGGTTTTTGTGTTTCAAACTAAGATGATTAATTCGGTTTTG 1756
 QY 2727 ATCTTTGCGAGAGCTGAGAAAGTGCAGACTTCAATGAGTTTCAAGAGAAAGATGA 2786
 Db 1757 ATCTTTGCGAGAGCTGAGAAAGTGCAGACTTCAATGAGTTTCAAGAGAAAGATGA 1806
 QY 2787 GGAGTTTGTGAAAGAGAGAGATGCTGATTAAGATCAAGAGAAAGATGAGAGACT 2846
 Db 1807 GGAGTTTGTGAAAGAGAGAGATGCTGATTAAGATCAAGAGAAAGATGAGAGACT 1866
 QY 2847 GAAAGAGAGCATCAAGAGAGATTTTGTATCTGGAGAAAGATTTGATGAGGCTTTGA 2906
 Db 1867 GAAAGAGAGCATCAAGAGAGATTTTGTATCTGGAGAAAGATTTGATGAGGCTTTGA 1926
 QY 2907 ACAGCTCATGTAACAGATGAGCTTCAACATGAGATGATTAAGACAAAGTCTGTACA 2966
 Db 1927 ACAGCTCATGTAACAGATGAGCTTCAACATGAGATGATTAAGACAAAGTCTGTACA 1986
 QY 2967 CAAGACAACTAAGTTTCTTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 3026
 Db 1987 CAAGACAACTAAGTTTCTTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTG 2046
 QY 3027 AGACTCCATTTAATCTAGACAAATCTAAGAGATTTATGATTAATTTCCCTCAATT 3086
 Db 2047 AGACTCCATTTAATCTAGACAAATCTAAGAGATTTATGATTAATTTCCCTCAATT 2106
 QY 3087 TTAGTAGAGGATCTAAGAGACATTAAGTTCTTGTGATTAACCA 3134
 Db 2107 TTAGTAGAGGATCTAAGAGACATTAAGTTCTTGTGATTAACCA 2154

RESULT 5
 BT004380
 LOCUS 1909 bp mRNA linear PLN 14-FEB-2003
 DEFINITION Arabidopsis thaliana clone U20243 unknown protein (At5g23570) mRNA,
 complete cds.
 ACCESSION BT004380
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 REFERENCE 1 (bases 1 to 1909)
 AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C.,
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamitani, A., Kawai, J., Kim, C.J., Narusaka, M.,
 Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P.,
 Southwick, A., Tzipi, M.G., Wu, T., Shinzaki, K., Davis, R.W.,
 Ecker, J.R. and Theologis, A.
 Arabidopsis Open Reading Frame (ORF) Clones
 Unpublished
 2 (bases 1 to 1909)
 Yamaoka, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C., Toriumi, M., Wong, C., Wu, H.C.,
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,

TITLE
JOURNAL
COMMENT

Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,
Nguyen, M., Palm, C.J., Sakurai, T., Saitou, M., Seki, M., Shim, P.,
Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W.,
Ecker, J.R. and Theologis, A.
Submitted (14-FEB-2003) Plant Gene Expression Center, 800 Buchanan
Street, Albany, CA 94710, USA
The RIKEN Genomic Sciences Center (GSC) members carried out the
collection and clustering of RFL CDNA (RFL CDNA : 'RIKEN
Arabidopsis Full-Length CDNA'): Seki, M., Narusaka, M., Ishida, J.,
Saitou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
Hayashizaki, Y. and Theologis, A.

The Salk, Stanford, PGEC (SSP) Consortium members constructed and
sequenced the PUNI (ORF) clones using the RFL CDNA: Yamada, K.,
Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M.,
Onodera, C.S., Quach, H.L., Tang, C.C., Tortum, M., Wong, C., Wu, H.C.,
Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,
Palm, C.J., Shim, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,
Ecker, J.R. and Theologis, A.

Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally
to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP
/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome
submitted to Genbank.

FEATURES

source
Location/Qualifiers
1. 1909
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="U20243"
/ecotype="Columbia"
/note="This clone is in pUNI 51."
1. 1909
/gene="At5g23570"
1. 1878
/gene="At5g23570"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="RA042374.1"
/db_xref="GI:28393933"
/translation="WSSRAQPGSKKKNVQSGTPRYVQVQIAGRIASSDDGSEW
EVIISKNNKRNPTSGKTGWSONSNPRPAGGQOQKRGSNVSRGNVSGRGNRG
IQANISGRGRALSRKYDNFVAPPPVSRPRLBGMWQAGSAKHIAVQEPFDVDD
VDNAISEENDSDALDSDDDLDASDDVDSQKSGSRKQNKFKFPGSLDLSIEO
INPROMCHPACONGPAGATDYNLHPLHARTKARBYKRLAEVLKDLQKRG
ASITPCEITYGKQKGLGDEKDYEIWMPVITIMTRLDNDKMLGNGNOLLEYP
DKYELARHSYGPQGRHWSVILAFBSATGTYEARELRELAEMGLDRIANQKSM
PQGVQQLYGFPLATKQDLDFNCHSQKTKLEKLSYQDMVKEHQISEDQQLAY
FKKLSKONKHAVLESLIEISMEKLRTEKLRTEKLRTEKLRTEKLRTEKLRTE
MDSEIKOIHERRDAKEENFEMLOOQERKXVGOOQOINPSSNDCKRAEVSPIEF
OEKEMEFVEREMILINDQEKMKEDMKRHHBEIPLEKFEDEALBQMLKHLNED
D"

1879. 1909
/gene="At5g23570"

ORIGIN

3' UTR

Query Match 45.6%; Score 1493; DB 8; Length 1909;
Best Local Similarity 83.5%; Pred. No. 0; Mismatches 0; Indels 376; Gaps 4;
Matches 1909; Conservative 0; Mismatches 0; Indels 376; Gaps 4;

QY 696 ATGAGTTCTAGGCTGGTCTCAATGTCTAAGAAAAAGACGTTCAAGGTTGTTAAGCCT 755
DB 1 ATGAGTTCTAGGCTGGTCTCAATGTCTAAGAAAAAGACGTTCAAGGTTAATAGCCT 60
QY 756 GAGGTTGAACGTTGTTCAAGGTTTGGCAGGAGCAGACCTGGTTCTTCAAGAAGAT 815
DB 61 GAGGTTGAACGTTGTTCAAGGTTTGGCAGGAGCAGACCTGGTTCTTCAAGAAGAT 120

QY 816 GAGAGAGAGTGGAGGTCAATTTCCAGAGAGAACAGAAACCAAGAAACCTTCTGGA 875
DB 121 GAGAGAGAGTGGAGGTCAATTTCCAGAGAGAACAGAAACCAAGAAACCTTCTGGA 180
QY 876 AAAAATTGGGTTTCTCAAGATTCCAAATCCCTTCTAGACTTGGGTGCTCAGACCAAGG 935
DB 181 AAAAATTGGGTTTCTCAAGATTCCAAATCCCTTCTAGACTTGGGTGCTCAGACCAAGG 240
QY 996 AGAGGTAGCAAGTATCTGGAGAGAGAGAAACAATGTATCCGGAGAGGTAAACGCAATGT 995
DB 241 AGAGGTAGCAAGTATCTGGAGAGAGAGAAACAATGTATCCGGAGAGGTAAACGCAATGT 300
QY 996 CGGGGCAATTCAGCTAACATATCTGTCGGGAGAGAGGCTTGGACAGAAAGTATGATAC 1055
DB 301 CGGGGCAATTCAGCTAACATATCTGTCGGGAGAGAGGCTTGGACAGAAAGTATGATAC 360
QY 1056 AACTTTTGGCAACCCCACTGTATCTGCCCCCTCTTTGGAAGAGATGGAATTGGAG 1115
DB 361 AACTTTTGGCAACCCCACTGTATCTGCCCCCTCTTTGGAAGAGATGGAATTGGAG 420
QY 1116 GCAAGAGAGGTTTCTGCTCAGACACAGCTGACAGAGATTCTCTGACGTGAGAGATAT 1175
DB 421 GCAAGAGAGGTTTCTGCTCAGACACAGCTGACAGAGATTCTCTGACGTGAGAGATAT 480
QY 1176 GTGATTAATGCTTCTGAGAGAGAGATGATTCGATGCTTGGATGATTCGATGACGAC 1235
DB 481 GTGATTAATGCTTCTGAGAGAGAGATGATTCGATGCTTGGATGATTCGATGACGAC 540
QY 1236 CTTCGAGATGATGATTAATGATCTCGATGTGATGATTAAGACCCATGATCGAAGAG 1295
DB 541 CTTCGAGATGATGATTAATGATCTCGATGTGATGATTAAGACCCATGATCGAAGAG 600
QY 1296 AATAAGTGGTCAAAAAGTCTTTGGCAGCTTGGATGATGATGATGATGATGATGAT 1355
DB 601 AATAAGTGGTCAAAAAGTCTTTGGCAGCTTGGATGATGATGATGATGATGATGAT 660
QY 1356 GAACCAAGAGGAGTGGATGATTCGACCTGTGCAAGCGACCTGTCGATCGATGATG 1415
DB 661 GAACCAAGAGGAGTGGATGATTCGACCTGTGCAAGCGACCTGTCGATCGATGATG 720
QY 1416 TTTAAGCTGACCCCTCTAAGCTCATGAGAGACAAAAGAGCTAGGAGATTAAGCTC 1475
DB 721 TTTAAGCTGACCCCTCTAAGCTCATGAGAGACAAAAGAGCTAGGAGATTAAGCTC 780
QY 1476 CATTAAGATTTGGCTGAAGATTTTAAGAAAGATCTTACAGATGAGAGGCGCATCTGCA 1535
DB 781 CATTAAGATTTGGCTGAAGATTTTAAGAAAGATCTTACAGATGAGAGGCGCATCTGCA 840
QY 1536 CTTTGTGTGAGATTTATGGGAGTGGAGAGGCTTGGGTGAGATGAAAGATTTATGAA 1595
DB 841 CTTTGTGTGAGATTTATGGGAGTGGAGAGGCTTGGGTGAGATGAAAGATTTATGAA 900
QY 1596 ATTGCTGGCCCTCAATGCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 1655
DB 901 ATTGCTGGCCCTCAATGCTCATCATCATCATCATCATCATCATCATCATCATCATCAT 960
QY 1656 AAGGTGAATCTCTGCTTCTTACTTCTTAAATTTTCTTGAATCTTACTGATCTTA 1715
DB 961 A----- 961
QY 1716 GAATGTTAATGTTAGTGGCTGCGCATGGGCAACCAAGCTGCTGGAATCTTGACAA 1775
DB 962 -----AGTGGCTGGCATGGGCAACCAAGCTGCTGGAATCTTGACAA 1007
QY 1776 GTATGAGCTCTTAAAGCAACGCAATTCATATGCTCACAGGGCCATCTGCGATGAGTGT 1835
DB 1008 GTATGAGCTCTTAAAGCAACGCAATTCATATGCTCACAGGGCCATCTGCGATGAGTGT 1067
QY 1836 TCTGATGTTTGAAGAGAGTGCACCTGGCTATTTGAGAGCCGAACGCTTCACCGGAGATT 1895
DB 1068 TCTGATGTTTGAAGAGAGTGCACCTGGCTATTTGAGAGCCGAACGCTTCACCGGAGATT 1127
QY 1896 AGCTGAGATGGGTTTGAATGAATTCCTGGGCTCAGAAAGCCGATATGTTTTCTGGAAG 1955

QY 1056 AACTTTGGGACCCCACTGATCTGCCCCCTCTTTGGAAGAGATGATGAG 1115
 Db 361 AACTTTGGGACCCCACTGATCTGCCCCCTCTTTGGAAGAGATGATGAG 420
 QY 1116 GCAGAGAGAGATTTCTGCTGACGACACAGCTGTGAGAGATTTCTGACGTGAGATGAT 1175
 Db 421 GCAGAGAGAGATTTCTGCTGACGACACAGCTGTGAGAGATTTCTGACGTGAGATGAT 480
 QY 1176 GTGAGATATGCTTTGAGAGAGAGATGATTTCCGATGCTTTGATGATTTCTGATGACGAC 1235
 Db 481 GTGAGATATGCTTTGAGAGAGAGATGATTTCCGATGCTTTGATGATTTCTGATGACGAC 540
 QY 1236 CTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1295
 Db 541 CTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 1296 AATTAAGTGTCAAAAAGTTCTTTGGCAGCTTGATGATGATGATGATGATGATGATGATGAT 1355
 Db 601 AATTAAGTGTCAAAAAGTTCTTTGGCAGCTTGATGATGATGATGATGATGATGATGATGAT 660
 QY 1356 GAACCAAG 1415
 Db 661 GAACCAAG 720
 QY 1416 TATTAACCTGACCCCTCTACTAGCTCATGAGAGACAAAAGAGCTTAGGAGATTAAGCTC 1475
 Db 721 TATTAACCTGACCCCTCTACTAGCTCATGAGAGACAAAAGAGCTTAGGAGATTAAGCTC 780
 QY 1476 CATAGAGATTTGGCTGAAAGTTTGAAGAGATCTACAGATGAGAGAGAGAGAGAGAGAGAT 1535
 Db 781 CATAGAGATTTGGCTGAAAGTTTGAAGAGATCTACAGATGAGAGAGAGAGAGAGAGAGAT 840
 QY 1536 CTTTGTGTGAGATTTTATGAT 1595
 Db 841 CTTTGTGTGAGATTTTATGAT 900
 QY 1596 ATTGTCTGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1655
 Db 901 ATTGTCTGAGCTTCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 QY 1656 AAGGTGAATCTCTGCTTTACTCTTTAATTTTCTCTGATTTACTGATCTTA 1715
 Db 961 A----- 961
 QY 1716 GAATGTTACATTTAGTGTGCTGAT 1775
 Db 962 -----AGTGTCTGAT 1007
 QY 1776 GTATGAGAGCTTTGAT 1835
 Db 1008 GTATGAGAGCTTTGAT 1067
 QY 1836 TCTGATGTTTGAAGAT 1895
 Db 1068 TCTGATGTTTGAAGAT 1127
 QY 1896 AGCTGAGATGAT 1955
 Db 1128 AGCTGAGATGAT 1187
 QY 1956 TGTTCGCAACTGTATGAGCTTCTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2015
 Db 1188 TGTTCGCAACTGTATGAGCTTCTTGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1247
 QY 2016 CTCTCAAGATTTCTCTCCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 2075
 Db 1248 CTCTCA----- 1253
 QY 2076 TTAAGATTTTGTGTGCTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2135
 Db 1254 -----AGG 1256

QY 2136 CAAAAGAGCTGAATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2195
 Db 1257 CAAAAGAGCTGAATTTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1316
 QY 2196 GCAGATCTCTGAT 2255
 Db 1317 GCAGATCTCTGAT 1376
 QY 2256 CAAGCAGCCAGAT 2315
 Db 1377 CAAGCAGCCAGAT 1436
 QY 2316 TGCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2375
 Db 1437 TGCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
 QY 2376 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2435
 Db 1497 AG----- 1498
 QY 2436 CATTTTGTATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2495
 Db 1499 -----AGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1514
 QY 2496 CAGGTTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2555
 Db 1515 CAGGTTTTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1574
 QY 2556 TTTGAGATGTTTGAAGAT 2615
 Db 1575 TTTGAGATGTTTGAAGAT 1634
 QY 2616 TAATCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2675
 Db 1635 TAATCCCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1665
 QY 2676 CTGAGGTTTTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2735
 Db 1666 ----- 1665
 QY 2736 CAGAGCTGAT 2795
 Db 1666 -AGAGCTGAT 1724
 QY 2796 GGAAG 2855
 Db 1725 GGAAG 1784
 QY 2856 GCATCAGAT 2915
 Db 1785 GCATCAGAT 1844
 QY 2916 GTTACAGAT 2949
 Db 1845 GTTACAGAT 1878
 RESULT 7
 ATHS28171 650 bp DNA linear PIN 29-MAR-2003
 LOCUS
 DEFINITION Arabidopsis thaliana T-DNA flanking sequence, left border, clone 157C08.
 ACCESSION AJ528171
 VERSION AJ528171.1 GI:26796431
 KEYWORDS left border; T-DNA flanking sequence.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Brakaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
 REFERENCE 1
 Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F.,
 Chauvin, S., Bechtold, N., Cruaud, C., DeRose, R., Pelletier, G.,

lepiniec, L., Caboche, M. and Lecharny, A.
T-DNA integration into the Arabidopsis genome depends on sequences
of pre-insertion sites

JOURNAL
EMBO Rep. 3 (12), 1152-1157 (2002)

MEDLINE
22363535

PUBMED
12446565

2 (bases 1 to 650)

REFERENCE
Balzerque, S.
Direct Submission
Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue
Gaston Cremieux, 91057 Evry cedex, FRANCE

COMMENT
PCR was performed on DNA from transformants of Arabidopsis thaliana
plants from INRA (Versailles). The DNA fragment(s) resulting from
the PCR were directly sequenced from the left or the right border
to determine the genomic sequence flanking the insertion. T-DNA
derived sequences were removed. Information to order the
corresponding mutant line and a link to a database providing a
graphical display of the insertion site are available at
<http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has
been generated in the framework of the French plant genomics
program 'Genoplante' (<http://www.genoplante.com> and
<http://genoplante-info.inrbio.gen.fr>).
Location/Qualifiers

FEATURES
source
1..650
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Massillaewskja"
/db_xref="taxon:3702"
/clone_idb="Arabidopsis thaliana T-DNA insertion lines"
1..650
/note="T-DNA flanking sequence
left border"

ORIGIN
Query Match 19.7%; Score 643.6; DB 8; Length 650;
Best Local Similarity 99.4%; Pred. No. 6.9e-131;
Matches 646; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1428 CCTCTCTAGCTATGCGAGGACAAAAGGAGCTTAAAGCTCCCTAAGAAATTG 1487
DB 1 CCTCTCTAGCTATGCGAGGACAAAAGGAGCTTAAAGCTCCCTAAGAAATTG 60

QY 1488 GCTGAAGTTTAAAGAGATCTACAGATGAGAGGGCATCTCTCTTGGTGAG 1547
DB 61 GCTGAAGTTTAAAGAGATCTACAGATGAGAGGGCATCTCTCTTGGTGAG 120

QY 1548 ATTATGCGAGTGAAGGGTTGGTGAAGATGAAAAGATTATGAAATTGCTGGCCT 1607
DB 121 ATTATGCGAGTGAAGGGTTGGTGAAGATGAAAAGATTATGAAATTGCTGGCCT 180

QY 1608 CCAATGCTATCTATCACTGAATTAAGCTGATGAGAGATACATTAAGTGAATTC 1667
DB 181 CCAATGCTATCTATCACTGAATTAAGCTGATGAGAGATACATTAAGTGAATTC 240

QY 1668 TTCTGCTTTACTTCTTTAATTTTCTTGATTTCTACTGATCTTAGAATTTCAAT 1727
DB 241 TTCTGCTTTACTTCTTTAATTTTCTTGATTTCTACTGATCTTAGAATTTCAAT 300

QY 1728 GTATGCTCGGAGTGGGCAACCAAGAGCTGTGGAATTAATTCGACAAGTATGAGCTCT 1787
DB 301 GTATGCTCGGAGTGGGCAACCAAGAGCTGTGGAATTAATTCGACAAGTATGAGCTCT 360

QY 1788 TAGAGACCGCATCTCTATGCTCAACAGGCGCATCTGGAGATAGTGTTCGATTTGA 1847
DB 361 TAGAGACCGCATCTCTATGCTCAACAGGCGCATCTGGAGATAGTGTTCGATTTGA 420

QY 1848 GAGCAGTGCACATGCGATTTGAGAGCGCAAGCGCTCCACCGGAGATTAGCTAGATGGG 1907
DB 421 GAGCAGTGCACATGCGATTTGAGAGCGCAAGCGCTCCACCGGAGATTAGCTAGATGGG 480

QY 1908 GTTAGATAGAAATTCCTGGGGTGAAGAGCGCATATGTTTTCTGAGAGTGTTCGCCAAT 1967
DB GTTAGATAGAAATTCCTGGGGTGAAGAGCGCATATGTTTTCTGAGAGTGTTCGCCAAT 1967

DB 481 GTTAGATAGAAATTCCTGGGGTGAAGAGCGCATATGTTTTCTGAGAGTGTTCGCCAAT 540

QY 1968 GTATGCTCTCTTGGCAAGACAGATGCTGAGCATATTCATCACTCAAGTTC 2027
DB 541 GTATGCTCTCTTGGCAAGACAGATGCTGAGCATATTCATCACTCAAGTTC 600

QY 2028 TCTCCCCCAAGAAATTTGATATATGCTTTAGTTTGTATGTAATTT 2077
DB 601 TCTCCCCCAAGAAATTTGATATATGCTTTAGTTTGTATGTAATTT 650

RESULT 8
AK064217
LOCUS
DEFINITION
Oryza sativa (japonica cultivar-group) cDNA clone:002-104-F07, full
insert sequence.
1556 bp mRNA linear JUN-24-JUL-2003

ACCESSION
AK064217 GI:32974235
VERSION
FLI CDNA; oligo-capping.
KEYWORDS
Oryza sativa (japonica cultivar-group)
SOURCE
Oryza sativa (japonica cultivar-group)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Barnardoideae; Oryzaceae; Oryza.

REFERENCE
AUTHORS
The Rice Full-length cDNA Consortium, National Institute of
Agrobiological Sciences Rice Full-length cDNA Project Team;
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Ootomo, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, W., Xie, Q., Lu, M.,
Naitkawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikura, J.,
Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, M., Matsubara, K., RIKEN;
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Itoh, M.,
Kagawa, T., Kondo, S., Komino, H., Miyazaki, A., Otsu, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shingawa, A., Shiraki, T.,
Yoshino, M. and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice
Science 301 (5631), 376-379 (2003)

TITLE
JOURNAL
MEDLINE
PUBMED
12869764
2 (bases 1 to 1556)

REFERENCE
AUTHORS
Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,
Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, K.,
Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,
Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K.,
Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,
Kawagashi, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,
Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Komino, H., Kouda, M.,
Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,
Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A.,
Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
Namiki, T., Naitkawa, R., Nikura, J., Nishi, K., Nomura, K.,
Numesaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Ooka, H.,
Otsu, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,
Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
Shingawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,
Tagawa, A., Takahashi, F., Takaku-Akashiri, S., Tanaka, T., Tomaru, A.,
Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,
Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
Yoshimura, A.
Direct Submission
Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
Agrobiological Sciences, Department of Molecular Genetics, Head of
Laboratory of Gene Expression; 2-1-2 Kamondai, Tsukuba, Ibaraki

Fujimura, T., Fukuda, S., Hanagasaki, T., Hara, A., Hashizume, M.,
 Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hizamoto, K., Hirooka, T.,
 Hori, F., Hotta, T., Ikeda, Y., Ikeda, R., Imamura, K., Hirooka, T.,
 Imotani, K., Ishibiki, Y., Ishii, Y., Ishikawa, M., Itoh, M., Kagaawa, I.,
 Kanagawa, S., Katoch, H., Kawagashita, N., Kawai, J., Kawamata, M.,
 Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M.,
 Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Konda, M.,
 Koya, S., Kurihara, C., Kuroseki, T., Kusumegi, T., Li, C., Lu, M.,
 Maeda, H., Matsubara, K., Matsuyama, T., Miura, U., Miyazaki, A.,
 Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,
 Namiki, T., Narikawa, R., Nilkura, J., Nishi, K., Nomura, K.,
 Numaneki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,
 Oeato, N., Ota, Y., Otomo, Y., Ryu, R., Satoh, H., Sakai, C., Sakai, K.,
 Sakazume, N., Sato, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,
 Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,
 Sugiyama, K., Suzuki, K., Suzuki, Y., Tagami, M., Tagami, Takeda, Y.,
 Tagawa, A., Takahashi, F., Takaku-Ikashira, S., Tanaka, T., Tomaru, A.,
 Toyota, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yanagi, M.,
 Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and
 Yoshimura, A.
 Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of
 Agricultural Sciences, Department of Molecular Genetics, Head of
 Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki
 305-8602, Japan (E-mail:skikuchi@nias.affrc.go.jp,
 Tel:81-29-838-7007, Fax:81-29-838-7007)
 This file is one of the 28K full-length cDNA clones from japonica


```

exon      /number=13
           complement(24024..24102)
           /gene="T14E10_10"
intron    /number=14
           complement(24103..24708)
           /gene="T14E10_10"
exon      /number=14
           complement(24709..24855)
           /gene="T14E10_10"
intron    /number=15
           complement(24856..25105)
           /gene="T14E10_10"
exon      /number=15
           complement(25106..25345)
           /gene="T14E10_10"
intron    /number=16
           complement(25346..25617)
           /gene="T14E10_10"
exon      /number=16
           complement(25618..25747)
           /gene="T14E10_10"
intron    /number=17
           complement(25748..25917)
           /gene="T14E10_20"
CDS       /gene="T14E10_20"
           join(27262..27331,27416..27633,27731..27982,28309..28586,
           29667..30057,30207..30665)
           /gene="T14E10_20"
           /note="Strong similarity to LeOPT1 oligopeptide
           transporter, Lycopersicon esculentum, EMBL:AF016713
           contains PTR2 family proton/oligopeptide symporters
           signatures AA73-97"
           /codon_start=1
           /product="oligopeptide transporter-like protein"
           /protein_id="CAB7565.1"
           /db_xref="GI:7258348"
           /db_xref="GOA:Q9M112"
           /translation="MADSTSLINKRTKGGNNALFIVVEIAERFAFGASNLITPEL
           TNEIGOSTATKAKNTWTGSCVSMFPIGAFADSLIGFKYLLTSPFLIGIYWL
           LSVTVARRRREKVFPMALYVAVGEGHKPCMTNADQGEANAEKAATSPFN
           WMAIVLASSIVLALIFIOERWSLIGSIIAGSVIAIVLIFLGIPIYRKQVPGS
           PFTRAVQVVAALKKWRLSTRHRYGLCYEEDEHLESTNSQVYLARTNQFRLD"

Query Match      5.6%; Score 181.8; DB 8; Length 94911;
Best Local Similarity 74.8%; Pred. No.2.8e-29;
Matches 291; Conservative 0; Mismatches 72; Indels 26; Gaps 4;

QY 2045 TGATATATGCTTTAGTTTGTCTGATGGAATTTAAAGTTTGTGGTCCGTGTTAATGCA 2104
      |||||
DB 57198 TCATTAACATTTCTAGCTTCCTCAATGGGATTTAGAGTTTGTGGTCTGTGGAATGCA 57139
      |||||
QY 2105 TCTGTTATGTTATCTATGATTCATTAGGCAAAACAGGCTGAAATTCGATGGAAT 2164
      |||||
DB 57138 TCA--TATGTATTTATGTATTTCTCATTTAGGCCAAACAGGCTGGAATTTGAGATGAAAT 57081
      |||||
QY 2165 CATACCAAGAGTGTGTTAAAGAGCTGAGCAGATCTCTGAGCAATCAGCAGCTGA 2224
      |||||
DB 57080 CTTACCAAGAAATGTTGTTAACTAGTGAAGCAGATCTCTGAGCAATCAGCAGCTGA 57021
      |||||
QY 2225 ACTACTTTAAGAACAGCTCTCAAAAAGAACAGCAGCAGCAAGCTGCTTGAGGAATCTC 2284
      |||||
DB 57020 TCTGGTTACAGAACAGATCACAACAACTGAAAG-----AGTATGGCTCTA 56974
      |||||
QY 2285 TGGAAATTTAGCGGAGAAAGCTCGTGAAGCTGCAAGATTAATCGGATCGTGAAGCA 2344
      |||||
DB 56973 TGGAAATTTAGTGAAGTCTGCGTGAAGCTGCAAGATTAATCGGATCGTGAAGCA 56919
      |||||
QY 2345 GAACATAGTGCAGATGAACAGACAGGAGAGAGGATGATTTTCC-----TAGAA 2398
      |||||
DB 56918 GAACAAAGATTCACATGAACAGACAGGAGAGGATGATTTTCCGGAAGAAAAA 56855
      |||||
QY 2399 ATCACAACTGACATTTGTATTAACCTA 2427
      |||||

```

```

DB 56858 CATACAACTGACATTTCTGATTAACCTA 56830

RESULT 11
CNS08CAO/c
LOCUS
DEFINITION
  CNS08CAO 91053 bp DNA linear PLN 21-NOV-2003
  Oryza sativa chromosome 12, . BAC OJ1561 A05 of library Monsanto
  from chromosome 12 of cultivar Nipponbare of sep. japonica of Oryza
  sativa (rice), complete sequence.
ACCESSION
  AL831811
  AL831811.4 GI:28892661
VERSION
  HTG.
KEYWORDS
  Oryza sativa (japonica cultivar-group)
SOURCE
  Oryza sativa (japonica cultivar-group)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
  Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
  1 (bases 1 to 91053)
    Choisne,N., Orjeda,G., Catolico,L., Demange,N., Wincker,P.,
    Segurens,B., Pelletier,E., Scarpelli,C., Salanoubat,M.,
    Weissenbach,J. and Quetier,F.
    Oryza sativa chromosome 12 sequencing
  2 (bases 1 to 91053)
    Unpublished
  Genoscope.
  Direct Submission
  Submitted (20-NOV-2003) Genoscope - Centre National de Sequencage :
  BP 191 91006 EVRY cedex - FRANCE (E-mail : sequef@genoscope.cns.fr
  - Web : www.genoscope.cns.fr)
  On Mar 9, 2003 this sequence version replaced gi:28460590.
  Center: Genoscope / Centre National de Sequencage
  Center code: GS
  Web site: http://www.genoscope.cns.fr/
  Contact: sequef@genoscope.cns.fr

-----
The following sequence is oriented from the T7 to the SP6 end. The
nucleotide sequence of this BAC clone was generated by combining
Monsanto, Syngenta and Genoscope sequencing data.
Upstream BAC (overlapping the T7 end) : OSJNBa0029N15 (AC-BX536967)
Downstream BAC (overlapping the SP6 end) : OJ1102_B11 (AC=AL713901)
-----
FINISHED SEGMENT STARTS AT BASE 1
FINISHED SEGMENT ENDS AT BASE 75085
-----

FEATURES
    source          Location/Qualifiers
                    1..91053
                    /organism="Oryza sativa (japonica cultivar-group)"
                    /mol_type="genomic DNA"
                    /cultivar="Nipponbare"
                    /sub_species="japonica"
                    /db_xref="taxon:39947"
                    /chromosome="12"
                    /clone="OJ1561 A05"
                    /clone_1fb="Monsanto"

ORIGIN
Query Match      5.3%; Score 174; DB 8; Length 91053;
Best Local Similarity 60.8%; Pred. No.1.4e-27;
Matches 320; Conservative 0; Mismatches 200; Indels 6; Gaps 2;

QY 1155 TTTCCTGACGTGGAGATGATGTGATATGCTTCTGAGGAGAGAAATGATCCAGTCT 1214
      |||||
DB 81521 TTGTCTGTTGATGATCCGAAACAGACAACTGTAAGGTGTAAGACATCGGACGATGAT 81462
      |||||
QY 1215 TTGATGATTTGATGACGACCTTGCAAGTATGATATGACTCGGATGTGATCAAAAG 1274
      |||||
DB 81461 AATGATGATGATATGATGACATTTGAGTGTGATGATTTGATGATGATGATGATAA 81402
      |||||
QY 1275 AGCCATGATCCGAAAGCAATTAAGTGTTCACAAAAGCTTTTGGCAGCTTGGATAGC 1334
      |||||
DB 81401 AGTTTGGAGACTCGGAAAAAATACAAAGTTTTCACAAAGTTTCTTGAAGTCTCTGAAGGCG 81342
      |||||
QY 1335 TTGTGATGACGACATTAATGAACCAAGAGGAGTGGCATTTGCCAGCTTGCAGAAC 1394
      |||||

```


TITLE

Kagawa, I., Kondo, S., Konno, H., Miyazaki, A., Osato, N., Ota, Y., Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from

JOURNAL

Science 301 (5631), 376-379 (2003)

MEDLINE

22752273

PUBMED

12869764

AUTHORS

2 (bases 1 to 2570)
Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Dol, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imanura, K., Imocani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Narikawa, R., Nikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toy, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, M., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and Yoshimura, A.

TITLE

Submitted (27-AUG-2002) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression, 2-1-2 Kamondai, Tsunoda, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tel:81-29-838-7007, Fax:81-29-838-7007)

COMMENT

This clone is one of the 28k full-length cDNA clones from japonica rice.
URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Dol, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Narikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K., and Murakami, K.
Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arikawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imanura, K., Imocani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toy, T., Waki, K., Yasunishi, A., and Hayashizaki, Y.

FEATURES

source

1.2570
Location/Qualifiers

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultiar="Nipponbare"
/db_xref="taxon:39947"
/clone="U023114M02"

ORIGIN

Query Match 4.8%; Score 155.8; DB 8; Length 2570;
Best Local Similarity 54.3%; Pred. No. 1,7e-23;
Matches 460; Conservative 0; Mismatches 307; Indels 80; Gaps 4;

Qy	1172	TGATGTGATTAATGCTTCTGAGAGAGAGATATCCGATGCTTTGGATGATTCGATGA	1231
Db	713	TGAATATGCAATGATGATGATGATTAATGATTCATATGACATGATTCGATGATGATGAT	772
Qy	1233	CGACCTTGCAATGATGATGATTAATGATCCGATGATGATGATGATGATGATGATGATGAT	1291
Db	773	TGATGATACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	832
Qy	1292	GCAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1351
Db	833	AAGGACAAAGCTGTTGAAAGATTTTGAACCTTCGAGGAAGTTAAGTTGAAACAGAT	892
Qy	1352	AAATGACCAAG	1411
Db	893	AAATGAAAAAGACTAAGCAATGATGATGATGATGATGATGATGATGATGATGATGATGAT	951
Qy	1412	TTGATTAATA---CTGCAACCTCTACTAGCTGATGATGATGATGATGATGATGATGATGAT	1468
Db	952	CTGATTAACAAAGGATTAACAGCTTTGATGATGATGATGATGATGATGATGATGATGATGAT	1011
Qy	1469	TGAGTTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1528
Db	1012	TGAGTTCATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1071
Qy	1529	TGTCATTCCTTGTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1588
Db	1072	AATGCAATATCTGCGAGCTTCTTGAATCTGAAAGATGCGAG---AAACATGTA	1128
Qy	1589	TTATGAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1648
Db	1129	TGCTCTTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1188
Qy	1649	TACCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1708
Db	1189	GGATGGA-----AGTGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGAT	1196
Qy	1709	GATCTTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1768
Db	1197	-----AGTGAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1235
Qy	1769	TGCAATTAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1828
Db	1236	TGATTAATTAATCACTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1295
Qy	1829	TGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1888
Db	1296	TGAGTGTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1355
Qy	1889	GGAGATTAAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1948
Db	1356	ATCACCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1415
Qy	1949	CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	2008
Db	1416	CTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT	1475
Qy	2009	ATCAACA 2015	
Db	1476	ATGACA 1482	

RESULT 14
AF542974 1564 bp mRNA linear PLN 02-JUL-2003
LOCUS AF542974
DEFINITION Triticum aestivum Emri mRNA, complete cds.
ACCESSION AF542974
VERSION AF542974.1 GI:32401385
KEYWORDS Triticum aestivum (bread wheat)
SOURCE

ORGANISM *Triticum aestivum*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.

REFERENCE
AUTHORS 1 (bases 1 to 1564)
TITLE Zhao, X., Li, Q. and Zhang, X.
JOURNAL Isolation and expression of a new kind of gene involve in
embryogenesis in *Triticum aestivum* L.
UNPUBLISHED
2 (bases 1 to 1564)
AUTHORS Zhao, X., Li, Q. and Zhang, X.
REFERENCE Direct Submission
TITLE Submitted (03-SEP-2002) College of Life Sciences, Shandong
JOURNAL Agricultural University, Daizong Street, Tai'an, Shandong 271018,
P.R. China

FEATURES
source Location/Qualifiers
1..1564
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="PH 82-2-2"
/db_xref="taxon:4565"
/dev_stage="embryo"
309..1223
/note="Involved in embryogenesis"
/codon_start=1
/product="Emr1"
/protein_id="AAP80862.1"
/db_xref="GI:32401386"
/translation="MSGNOELLEYSDYATKARHAYGPGHGMVLIPESSAVGME
AERLHKHFDORTDTONRVRVFLPGKRLQYGLARKEDEMEFPHNCOGSRKY
EMRSHNEVVAQKMSQEDNOQLYLNKRVKTEORSKVEETLGLVITOKLAEEM
IIVRSKAEKSEEESEKSEFPHQINIHAEDEKSEFERLQERAKARCCD
VDSCTENRRLRKEOVORFICOVDOVEFAERDEMIKAEKVKQLKKEVMAKVE
LEKFDALITGLMEKHRTGTQASSSP"

CDS
ORIGIN

Query Match 3.7%; Score 120.8; DB 8; Length 1564;
Best Local Similarity 61.2%; Pred. No. 8.8e-16;
Matches 213; Conservative 0; Mismatches 132; Indels 3; Gaps 1;

1717 AATGTTACATTTGATGCTGCTGGCAGCAACCAAGAGCTGCTGATCTTGACACAG 1776
285 ATGATGCTCAGTCAAGTGTAGGCAATGAGCAACCAAGAACTTCTGATATTTCAAGTAT 344
1777 TATGAGGCTCTTGAAGCAGCCATTCCTATGCTCAAGGCCATCTGGAGATGATGTT 1836
345 TATGCTGCAACCAAGCAGCCATGATGCTCAGGTGGCAGCGGTGATGATGTT 404
1837 CTGATGTTTGAAGCAGCTGCTATTTGGAGCCGCAAGCCCTCCACCGGAGTTA 1896
405 CTAAATATTTGAAGCTGCTGCTGGGCTATATGAGGCAAGACGTTCTTCAATGCAATTT 464
1897 GCTGAGATGGGGTTAGATGAATTTGCTG---GGGTCAAGAGCGCATGTTTCTGGA 1953
465 ATTGATCAAGAGACAGACAGGACACTTGGCAGAAATCGCAGGGTTCCTTTTAACTGAT 524
1954 GGTGCTGCGCAACTGATGCTGCTTCTGCAACGAGAGAGATGAGCATTTCAATCA 2013
525 GGGAGAGACAAATATAGGTTTCTTGGCAGAAAGAGAGACATGAGACTTTCAATAGA 584
2014 CACTCTCAAGGTTCTCTCCCAAGAAATTTGATATATGCTTTTACT 2061
585 CACTGCCAAGGAAAGCCGCTGAATATACAGATGAGATCTCATAT 632

RESULT 15
AF469493 513 bp mRNA linear PLN 02-JUL-2003
LOCUS *Triticum aestivum* H0TR mRNA, partial cde.
DEFINITION
ACCESSION AF469493
VERSION AF469493.1 GI:32400755
KEYWORDS
SOURCE *Triticum aestivum* (bread wheat)

ORGANISM *Triticum aestivum*
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.

REFERENCE
AUTHORS 1 (bases 1 to 513)
TITLE Li, J.R., Wang, F., Li, Q.Z. and Zhang, X.S.
JOURNAL Gene isolation and expression of a new Zn-finger
UNPUBLISHED
2 (bases 1 to 513)
AUTHORS Li, J.R., Wang, F., Li, Q.Z. and Zhang, X.S.
REFERENCE Direct Submission
TITLE Submitted (18-JAN-2002) College of Life Science, Shandong
JOURNAL Agricultural University, Dai zong Street 61, Taidan, Shandong
271018, P.R. China

FEATURES
source Location/Qualifiers
1..513
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="PH 82-2-2"
/db_xref="taxon:4565"
/dev_stage="seeds"
1..513
/note="Similar to H0TR of *Hordeum vulgare*"
/codon_start=1
/product="H0TR"
/protein_id="AAP80610.1"
/db_xref="GI:32400756"
/translation="WLMGNOELLEYSDYATKARHAYGPGHGMVLIPESSAVG
YMAERLHKHFDORTDTONRVRVFLPGKRLQYGLARKEDEMEFPHNCOGSR
IKYEMSHNEVVAQKMSQEDNOQLYLNKRVKTEORSKVEETLGLVITOKLAEEM
ENIFVRSKA"

CDS
ORIGIN

Query Match 3.6%; Score 117; DB 8; Length 513;
Best Local Similarity 61.6%; Pred. No. 6.3e-15;
Matches 205; Conservative 0; Mismatches 125; Indels 3; Gaps 1;

1732 TGGCTGGCAGTGGCAACCAAGAGCTGCTGAAATCTTGACAAAGATGAGGCTTTAGA 1791
1 TGGTTAGCAGTGGGCAACCAAGAACTTCTTGATATTTCAAGTATGCTGGACCAAA 60
1792 GCAGCCATTCCTATGATGCCAAGGCCATCTGGATGAGTCTTGATGTTTGAAGC 1851
61 GCAGCCATTCCTATGATGCCAAGGCCATCTGGATGAGTCTTGATGTTTGAAGC 120
1852 AGTGCATCTGCTATTTGAGGCGCAAGCCCTCCACCGGAGTTAGCTGAGTGGGTTA 1911
121 TGTGCTGTGGCTATATGAGGCAAGCGTCTTCAATGACATTTTATGATCAAGAACA 180
1912 GATGAATTTGCTG---GGGTCAAGAGCGCATGATGTTTCTGAGAGTGTGCAACTG 1968
181 GACAGGACACTTGGCAGAAATCGCAGGGTTCCTTTCTTACCTGTGGAGAGACAAATTA 240
1969 TATGCTTCTTGGCAACGAAGCATCTGACATTTCAATCAACTCTCAAGTTCT 2028
241 TAGGTTTCTTACCCAGAAAGAGACATGAGACTTTCAATAGACATGCGCAAGGAAA 300
2029 CTCGCCAAGAAATTTGATATATGCTTTTACT 2061
301 AGCGCTGAAATTCAGAGATGAGATCTCATAT 333

Search completed: April 8, 2004, 13:33:35
Job time : 8568.15 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 14:30:53 ; Search time 1814.04 Seconds
(without alignments)
16605.743 Million cell updates/sec

Title: US-10-030-829-1_COPY_1_695
Perfect score: 695
Sequence: 1 gacaaacaaacaaatcaa.....ggcgcctatagggtgtaaa 695

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenBml: *
1: gb_da: *
2: gb_hcg: *
3: gb_in: *
4: gb_ov: *
5: gb_ov: *
6: gb_ov: *
7: gb_ov: *
8: gb_ov: *
9: gb_ov: *
10: gb_ov: *
11: gb_ov: *
12: gb_ov: *
13: gb_ov: *
14: gb_ov: *
15: gb_ov: *
16: gb_ov: *
17: gb_ov: *
18: gb_ov: *
19: gb_ov: *
20: gb_ov: *
21: gb_ov: *
22: gb_ov: *
23: gb_ov: *
24: gb_ov: *
25: gb_ov: *
26: gb_ov: *
27: gb_ov: *
28: gb_ov: *
29: gb_ov: *
30: gb_ov: *
31: gb_ov: *
32: gb_ov: *
33: gb_ov: *
34: gb_ov: *
35: gb_ov: *
36: gb_ov: *
37: gb_ov: *
38: gb_ov: *
39: gb_ov: *
40: gb_ov: *
41: gb_ov: *

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	695	100.0	3275	6	AX078760
2	695	100.0	81365	8	AB025633
3	85.6	12.3	7218	6	166494
4	84	12.1	2162	8	BT002944
5	83.4	12.0	184535	2	CNS057CJ
6	81.2	11.7	349980	6	AX344566
7	72.8	10.5	1407	8	AJ592026
8	70	10.1	1453	8	AJ591978
9	70	10.1	66993	2	AC138074
10	69.6	10.0	1434	8	AJ592058
11	69.2	10.0	3364	14	TSP418778
12	69	9.9	88932	2	AC022680
13	68.8	9.9	6668	6	AX346599
14	68.8	9.9	81120	2	AC022851
15	68.4	9.8	810	8	AJ592180
16	68.2	9.8	56857	2	AC021917
17	68	9.8	1184	8	AJ591982
18	67.6	9.7	151368	2	AC142664
19	67.4	9.7	156550	2	AC015830
20	67.4	9.7	196853	9	CNS01BPD
21	67	9.6	1228	8	AJ592059
22	67	9.6	60565	2	AC023852
23	67	9.6	127709	2	CNS057C1
24	67	9.6	202083	2	AC023833
25	66.8	9.6	164288	2	AC128949
26	66.6	9.6	749	9	HSN335745
27	66.6	9.6	9539	6	AX277889
28	66.6	9.6	9539	6	AX323566
29	66.4	9.6	83440	2	AC024285
30	66.4	9.6	194180	2	AC069140
31	66	9.5	6644	6	E23355
32	66	9.5	7372	6	E23357
33	66	9.5	7797	6	E23355
34	66	9.5	7996	6	E23359
35	66	9.5	14006	6	AX346860
36	66	9.5	56693	2	AC084077
37	66	9.5	143585	2	AC013349
38	65.8	9.5	64767	2	AC102701
39	65.8	9.5	273807	2	AC025421
40	65.8	9.5	289973	2	AC135678
41	65.6	9.4	113515	6	AX347076
42	65.4	9.4	6079	6	AX345323
43	65.4	9.4	93791	2	AC138073
44	65.4	9.4	174384	2	AC009524
45	65.4	9.4	242885	2	AC097836

ALIGNMENTS

RESULT 1
AX078760 3275 bp DNA linear PAT 22-FEB-2001
LOCUS Sequence 1 from Patent WO0105951.
DEFINITION AX078760
ACCESSION AX078760
VERSION AX078760.1 GI:1158379
KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
AUTHORS Beclin,C., Elmayer,T. and Vaucheret,H.
TITLE Novel sgsl plant gene and use thereof

JOURNAL Patent: WO 0105951-A 1 25-JAN-2001;
 AVENTIS CROPS SCIENCE S.A. (FR) ; INSTITUT NATIONAL DE LA RECHERCHE

AGRONOMIQUE (FR)

FEATURES
 source 1.3275
 Location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"
 /note="p356AD"
 primer_bind complement(2926..2952)
 primer_bind /note="p356Y"

ORIGIN

Query Match 100.0%; Score 695; DB 6; Length 3275;
 Best Local Similarity 100.0%; Pred. No. 1.2e-130; Indels 0; Gaps 0;
 Matches 695; Conservative 0; Mismatches 0;

1 GACAAACAAACAAATTTAGCAAGTCATGTCGATGATTAATTAATAGTGGACAA 60
 1 GACAAACAAACAAATTTAGCAAGTCATGTCGATGATTAATTAATAGTGGACAA 60
 1 TTAAGTTAAGCGAAGAAAGAAAGAAAGGTAAGAAATGAAACAAATCAACGAT 120
 1 TTAAGTTAAGCGAAGAAAGAAAGAAAGGTAAGAAATGAAACAAATCAACGAT 120
 61 TTAAGTTAAGCGAAGAAAGAAAGAAAGGTAAGAAATGAAACAAATCAACGAT 120
 61 TTAAGTTAAGCGAAGAAAGAAAGAAAGGTAAGAAATGAAACAAATCAACGAT 120
 121 GAAATTTGAGTCGACAAATCGAAGAAAGAGGCGCTTTAGAGCTTAATAGCTTC 180
 121 GAAATTTGAGTCGACAAATCGAAGAAAGAGGCGCTTTAGAGCTTAATAGCTTC 180
 121 GAAATTTGAGTCGACAAATCGAAGAAAGAGGCGCTTTAGAGCTTAATAGCTTC 180
 121 GAAATTTGAGTCGACAAATCGAAGAAAGAGGCGCTTTAGAGCTTAATAGCTTC 180
 181 ATTGTCCTC 240
 181 ATTGTCCTC 240
 181 ATTGTCCTC 240
 241 TCTCCGCGCTTTAACTTACGTCCTCCGCTTCTCTCTCTCTCTCTCTCTCTC 300
 241 TCTCCGCGCTTTAACTTACGTCCTCCGCTTCTCTCTCTCTCTCTCTCTCTC 300
 241 TCTCCGCGCTTTAACTTACGTCCTCCGCTTCTCTCTCTCTCTCTCTCTCTC 300
 301 CCTCGGATCGCTCAGCGCAGCATTCCTGCTGATTTCTTTTCTGCTGAGAA 360
 301 CCTCGGATCGCTCAGCGCAGCATTCCTGCTGATTTCTTTTCTGCTGAGAA 360
 301 CCTCGGATCGCTCAGCGCAGCATTCCTGCTGATTTCTTTTCTGCTGAGAA 360
 361 ATTGCTTAAGTTCTGATTTGAAAGTTTGGCTGAGAGGTTACTTTTCTCTATA 420
 361 ATTGCTTAAGTTCTGATTTGAAAGTTTGGCTGAGAGGTTACTTTTCTCTATA 420
 361 ATTGCTTAAGTTCTGATTTGAAAGTTTGGCTGAGAGGTTACTTTTCTCTATA 420
 421 TTTTATAGTTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 480
 421 TTTTATAGTTCTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 480
 481 TTACCGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 540
 481 TTACCGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 540
 481 TTACCGTTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACTTAACT 540
 541 TTGTTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 541 TTGTTTCAATTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 601 ATGCGTTTGGCTTGTATTAATCTTCAAGTTTGAATTTGATTTTGGTAGG 660
 601 ATGCGTTTGGCTTGTATTAATCTTCAAGTTTGAATTTGATTTTGGTAGG 660
 601 ATGCGTTTGGCTTGTATTAATCTTCAAGTTTGAATTTGATTTTGGTAGG 660
 661 ACTGTGGGTTTCTTGGTGGCTAAGGTTTAAAA 695
 661 ACTGTGGGTTTCTTGGTGGCTAAGGTTTAAAA 695

RESULT 2
 AB025633 81365 bp DNA linear PLN 27-DEC-2000
 LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone:WQMI.
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, Pl clone:WQMI.
 ACCESSION AB025633
 VERSION AB025633.2 GI:10178221

KEYWORDS

SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (sites)
 Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H.
 and Tabata, S.
 Structural analysis of Arabidopsis thaliana chromosome 5. X.
 Sequence features of the regions of 3,076,755 bp covered by sixty
 Pl and TAC clones
 DNA Res. 7 (1), 31-63 (2000)

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

JOURNAL

COMMENT

Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (sites)
 Sato, S., Nakamura, Y., Kaneko, T., Katoh, T., Asamizu, E., Kotani, H.
 and Tabata, S.
 Structural analysis of Arabidopsis thaliana chromosome 5. X.
 Sequence features of the regions of 3,076,755 bp covered by sixty
 Pl and TAC clones
 DNA Res. 7 (1), 31-63 (2000)
 On Sep 15, 2000 this sequence version replaced gi:4589439.
 Address for correspondence: kaos@kazusa.or.jp
 For the latest information on annotation of this clone, please see
 http://www.kazusa.or.jp/kaos/cgi-bin/sgd_graph.cgi?c=WQMI
 Genes with similarity to proteins in the databases are described in
 'product' or 'note' qualifiers. Genes that have no significant
 protein similarity are described as 'unknown protein'.
 The software programs used to predict genes include: Grail
 (Informatics Group, Oak Ridge National Laboratory,
 http://compbio.ornl.gov/Grail-1.3/),
 GENSCAN (Chris Burge, MIT, http://CCR-081.mit.edu/GENSCAN.html),
 NERGEN2 (S.M. Hebsgaard, et al., CBS, Technical University of
 Denmark, http://www.cbs.dtu.dk/services/NERGEN2/) and
 SplicePredictor (Volker Brendel, Stanford University,
 http://gremlim.zool.iastate.edu/cgi-bin/ep.cgi).
 Genes encoding tRNAs are predicted by tRNAscan-SE
 (Sean Eddy, Washington University School of Medicine, St. Louis,
 http://genome.wustl.edu/eddy/tRNAscan-SE/).
 This sequence may not be the entire insert of this clone. It may be
 shorter because we remove overlaps between neighboring submissions.
 The 5' clone is K19M13 and the 3' clone is M011.

FEATURES

source

CDS

CDS

1..81365
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="WQMI"
 /clone_1fb="M16su Pl"
 /clone_3184..3403,3474..3595,3681..3723,3777..3898,
 /join(3184..3403,3474..3595,3681..3723,3777..3898,
 /4042..4198,4445..4515,4601..4774,4888..4988,5300..5420)
 /note="gb|AA01580.1
 gene_id:WQMI.24
 similar to unknown protein"
 /codon_start=1
 /evidence=not experimental
 /protein_id="BA97251.1"
 /db_xref="GI:8809710"
 /translation="MTVPGTGVDSPPAHKVDTPPKVQPLRISTKTVDDAONPS
 YGAFDPSSNSPLSPVPEPSSSFGSGDDPLGLEDLQISGEYPPHLEACG
 YSINGTSCNFEVYIDESSRERTSNAGAKPNYLVADVDLYLAEV
 OPIDDRNRKPRAPKGVYDLTVLDVHTLLFSADCSVFPMFAFICVFLISG
 ACDDPMDQSNIEKTLTGAASVKSILAVGVDIWEATLISIRGEYSIKCSLDLIAE
 KPSASTVTIIPGQPAELVITISDGSSEHLRDNPSPLIGRDEIVLIRLFIKRW
 FFKALNSFYHFRCSKSRGNL"
 /join(6317..6428,6508..6629,6714..6756,6841..6932,
 /7121..7169,7396..7466,7550..7663,7785..7882,8112..8181)
 /note="gb|AA01580.1
 gene_id:WQMI.23
 similar to unknown protein"

/codon_start=1
/evidence=not experimental
/protein_id="BAA97250.1"
/db_xref="GI:8809709"
/translation="MOPGTSVYESPALNADDEPPISKVYVNDTKIIFPDGDDAPL
PALENIQISGEYPGHELOACGSIINGTSCNFEWCHLEDGSVNIDAKKENYLV
ADVQGLAIEVQPLDDNRKELVVFANDNRKLAICHEPMOSNIDKTLHTGASXV
SLAIGFVHIMEAATLSIEREGYTIKCNNDLTTEKESASTAVIKPEKPAEVLITGSD
GSEHCLRVNEMPDILCKEKKRLVLYVVFEOKL"
complement(join(8642..8705,8794..8863,8978..9387,
9584..10347))
/note="gene_id:MOM1.22
unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97249.1"
/db_xref="GI:8809708"
/translation="MSWMKGSQWTAFLDKORQKGLSEVEGDFPEPVSTSVNASF
GVGRRLRNHSESEKSSVLLPPSRFPALTEKDCNQBERGGCCRRKEDTLSPVNS
HDLAFTKLKENMSWADNLIIRDVLSTEDDEFEALFLKGVNCSGKEDEPTSLIGY
SSDNRSREYRTEKTVTSSVKAABSTPEADAGYDLNDSGSPVNAADNKEPPDI
SELDSTIQSLQSLPIEPWEEDDLVSHRKDALKWRKSNSRAQNAFORDHASA
KCHSDAREDMLAERKLAEEAAKLIIGITKNDNWKLDHGLHATEAVALQERLQ
IGHFTVNSVSBNRGRSNAALRSASQEPFLDEGHCQRTSRELRNLSQVITG
ICKHSRQASLPLAVKTFEDNRYRDETRPGVITVRPFRRHS"
complement(12417..13424)
/note="contains similarity to unknown protein
gb|AAE27018.1
gene_id:MOM1.21"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97248.1"
/db_xref="GI:8809707"
/translation="NATDSQPNOKLTPLKTRIALTVISTMTDNRQRPDCTINRRRL
LDFRAPNPKNVNIIVSTDSFVVDSDRLMFLYTHVSGDKIPVVFPHGGFAFLS
PAAPYDYNCRFRFARLPAVVISVNYRLAEPHRYPQDGDGPALEIENNGSLIPA
NADLSRCPFAGDSAGNIAHVAIRICRPRSFTLVKILGILSIQPFSGERTAE
KOLVGAIVSPRTDCKMAGLNDRBAGVPAVADISGLDYETPVVAVAGFPL
KQMRQSYVEMLKCGKATLIEVPMNFHAFYIPELPEAGQLIMRIKQVDERVASLS
A"
join(14922..14981,15072..15167,15563..15666,15755..15891,
16096..16188)
/note="gene_id:MOM1.20"
/codon_start=1
/evidence=not experimental
/product="50S ribosomal protein L24"
/protein_id="BAA97247.1"
/db_xref="GI:8809706"
/translation="MGMAAEKLIIRMKILRGDNWMIIRKDKGETGITIKRVISQR
VVEGKNLIKKHIIKGGPDHEGGIFVTEAPLHASNVQVDPVTRGCKVAVKLEDTX
VVAARCTGSGSIIIRPELIIKIRATPRPTAGKXDPMEFVMEQYDADATGGMPL"
join(16823..16952,17143..17320,17412..17493,17675..17851,
17943..18110,18199..18390)
/note="gene_id:MOM1.19"
/codon_start=1
/evidence=not experimental
/product="26S proteasome, non-ATPase regulatory subunit"
/protein_id="BAA97246.1"
/db_xref="GI:8809705"
/translation="MERLQIRFAGCGLGHASPDSPFLDTSQVYISLALIKMLKHC
RAGVPEVGLMGEFVDEYTVVDFVFMPSGTSVEAEDVHETLMLKOTG
REMVGVGSHSGEGCMISGVDINFGALNDAVAVVVDPIOSVAGKVVIAFR
SINPOTIMGCRPOTSNLGHKNPSIOLIHGNRHYSTAINRKNELEBKMLN
LHKKWTDLIRRFPTSKINQVQEMSLAKAKNAVQEBDELSPKLAIVANGR
QAKKLEHVSNSNSNIIVQTLGMLDVVF"
complement(join(18965..19043,19240..19298,19419..19536,
19666..19739,19908..19961,20387..20530))
/note="contains similarity to unknown protein
emb|CAA22897.1
gene_id:MOM1.18"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97245.1"

/db_xref="GI:8809704"
/translation="MDKNQALIEKMKLVGEVEDEQQADEESSLSFWMEDLRNCA
LTKORFVGAICLSAGICTLISMLVFNVPVFGTITLGNMAGSTAFILGPOROV
TMLADPARIVATALIASIIIALFCALVYRNKLLTLALLIEFTGLIWSLSIFPAR
TMVSKIRNFCPTFEF"
join(22672..22634,23708..23999,24111..24355,24458..24624,
24715..24925)
/note="emb|CA862356.1
gene_id:MOM1.17
similar to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97244.1"
/db_xref="GI:8809703"
/translation="MSSRAGPMSKEKNVQGVREPEVOLVQIAGTRLASODDGEW
EYISKNNKRGNTSGKTWVNSONSPPAMGCOOQGRGNSGRANNVSGNNGRG
IOANTSGRALSRTYDNNFVAPPPVSRPELEGNNMQARGSAOHTAVQEPDYED
VMAEERNDSDALDSDDDLSDDYSDVSQSHSGSRKQNMFKFPGSLDLSIEQ
INEPORHNPACONGPQAIIDVYNLHPILAHARTGARVXKLHRELALENDOLMRG
ASVIFGEIYQGMKGLGDEKDYELVMPPMVIMNTRLDKDNDKMLGONQELLEY
DKVELARHSGVQGHGMSVLMFESSATGYLRELRHRELAELGLOBIANGQKSM
PSGGRQLYGLPLATQDDIIPNOSHGRTRLEKLSQVQMYVKEIROSENOQNTY
FNKLSKQNRKAKVLESLEIWEKLRKTAEDNRIVRQRTKQHEONRENDADHREF
MDSIQIHERHRAKENEEMLOQDERAVVGOQONINPSNDDCKRAEEVSSIEF
QEKEMEEFVEEREMLIKQEKEMEDMKRHHHEIFDLKEFDEALEQIIMYKRGLHNE
D"
join(25614..25838,26061..26207,26324..26438,26516..26610,
26710..26813,27522..27765,27943..28036,28115..28305,
28385..28604,28688..28855,28944..29092)
/note="gb|AAC97420.1
gene_id:MOM1.16
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97243.1"
/db_xref="GI:8809702"
/translation="MAPPAGQAAVAEAGADGAPPOQGRGFTISGIVRAVWFY
ASKPSPKQKMPDPAQOLMNTLPHKESLDMMFVYLSQEFNDFGNDRALYMETN
IPYAVTPEISRTKSLTYPSRTLNONSIVAHIFPARGSPIDTDPDYOPLNSPR
THAATVPPKQKNNKKSILGSPKSDSBEPEVGVGKSDPKREVPREWISLKNP
VTINLVDPFQADLVEFTGNYVPTFYFNELKDFEFTVNSFTLPLNLEISF
ISMMKQOLFQOVDQFQORSYGMIDSEBELKVFLEGNFYLLGITMFVMSLVSF
DELFKNDIOFNKNNKSMGLSAKSVNLFIQFVIFLYLLDNDSMMILASSGVGC
IEFKIKGAMRIEVDKSGMIPRLRPHDESVNSNTEKYDDIAIKELSVLLLVYGL
SIVSLAVERKHSVMSIISLTSVYMGFTMGCQLPNTYKLSVAHLPMQOMTYKE
LNTITDDLPAPVITKPIIHRLSVFPDDVIFLYLYQRNVPYDKTRVNEFGGDEET
AEKRLITEKEEDKXTN"
complement(join(29439..29576,29661..29865,29985..30268,
30341..30493,30579..30722,30956..31484))
/note="gene_id:MOM1.15

Query Match 100.0%; Score 695; DB 8; Length 81365;
Best Local Similarity 100.0%; Pred. No. 7.8e-131;
Matches 695; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CDS

QY 1 GACAAACAAACAAATTAAGCAAGTCATGTCGTAAGCAATAATTAATAGTGGACAA 60
DB 21977 GACAAACAAACAAATTAAGCAAGTCATGTCGTAAGCAATAATTAATAGTGGACAA 22036
QY 61 TTAAGTTAAGCAAAAGCAAAAGCAAAAGCAAAATGAACAAATCAACTGAAT 120
DB 22037 TTAAGTTAAGCAAAAGCAAAAGCAAAAGCAAAATGAACAAATCAACTGAAT 22096
QY 121 GAAATTTGAGTCCGAATTCGAAACAGAGCCGCTTTTAAGCTTAATTAAGCTTCC 180
DB 22097 GAAATTTGAGTCCGAATTCGAAACAGAGCCGCTTTTAAGCTTAATTAAGCTTCC 22156
QY 181 ATTGATCTCTTCTGATGATTTATTTCTCTCGGAGTCTGATCACTACTCTCAC 240
DB 22157 ATTGATCTCTTCTGATGATTTATTTCTCTCGGAGTCTGATCACTACTCTCAC 22216
QY 241 TCTCGGCGCTTTAACTTACCTCTCGGCTTTACTCTGAATTTTCTGCTTAAG 300

Db 22217 TCTCGGCGCTTAACTAGCTTCCGTGCTTACTGTAAGTTTTCGCTTAAG 22276
 QY 301 CCTCGAGTGCCTCAACCGCATGCTTGTGCTGCAATTTCTTTTCTTGGCGGAAA 360
 Db 22277 CCTCGAGTGCCTCAACCGCATGCTTGTGCTGCAATTTCTTTTCTTGGCGGAAA 22336
 QY 361 ATTCGCTTAAATGTTCTGCAATTCGCAAGTTTGTGCTAAAGGCTTACTTTTCCCTA 420
 Db 22337 ATTCGCTTAAATGTTCTGCAATTCGCAAGTTTGTGCTAAAGGCTTACTTTTCCCTA 22396
 QY 421 TTTTAAAGTCTTAGGTAACGATACCTGCGCTTACTGTTTGTGTAATTTGTGTGCT 480
 Db 22397 TTTTAAAGTCTTAGGTAACGATACCTGCGCTTACTGTTTGTGTAATTTGTGTGCT 22456
 QY 481 TTTACGCTTAAATGTTCTGCAATTCGCAAGTTTGTGCTAAAGGCTTACTTTTCCCTA 540
 Db 22457 TTTACGCTTAAATGTTCTGCAATTCGCAAGTTTGTGCTAAAGGCTTACTTTTCCCTA 22516
 QY 541 TTTGTTCAATAAATGCTGATGATCTACTTTTGTGCTTGTGCTTGTGCTTGTGCT 600
 Db 22517 TTTGTTCAATAAATGCTGATGATCTACTTTTGTGCTTGTGCTTGTGCTTGTGCT 22576
 QY 601 ATGCGTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 660
 Db 22577 ATGCGTTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 22636
 QY 661 ACTGTGGGTTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 695
 Db 22637 ACTGTGGGTTTCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCTTGTGCT 22671

RESULT 3
 LOCUS 166494 7218 bp DNA linear PAT 28-DEC-1997
 DEFINITION Sequence 14 from patent US 5670367.
 ACCESSION 166494
 VERSION 166494.1 GI:2724471
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 7218)
 AUTHORS Dörner, F., Schefflinger, F. and Falkner, F. Gunter.
 TITLE Recombinant fowlpox virus
 JOURNAL Patent: US 5670367-A 14-23-SEP-1997;
 FEATURES
 source 1..7218
 /organism="unknown"
 /mol_type="unassigned DNA"

ORIGIN
 Query Match 12.3%; Score 85.6; DB 6; Length 7218;
 Best Local Similarity 6.4%; Pred. No. 1.7e-07;
 Matches 25; Conservative 232; Mismatches 131; Indels 0; Gaps 0;
 QY 174 CTTCCTATTTGTCTCTCTCTGCAATTTCTTCTCCGCGAGTCCGATCTACTA 233
 Db 1094 YY 1153
 QY 234 CTCTCACTCTCCGCGCTTAACTAGTCTCGCTGCTTACTGTAAGTTTCTGTC 293
 Db 1154 YY 1213
 QY 294 CTTAAGCTCCGATGCGCTCAACGATGATCTGCTGCTGATTTCTTTTCTCCGC 353
 Db 1214 YY 1273
 QY 354 TGAATAATGCGCGATGCTGATTTGCAAGTTTGTGCTTAAAGGCTTACTTTT 413
 Db 1274 YY 1333
 QY 414 CCTTATTTTAAATGTTCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 473

Db 1334 YY 1393
 QY 474 TTGCTTAAACGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 533
 Db 1394 YY 1453
 QY 534 TTGCTTAAACGCTTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 561
 Db 1454 CTTAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 1481

RESULT 4
 LOCUS BT002944 2162 bp mRNA linear PLN 15-JAN-2003
 DEFINITION Arabidopsis thaliana clone RAFL14-93-K05 (R20243) unknown protein
 (A5523570) mRNA, complete cds.
 ACCESSION BT002944
 VERSION BT002944.1 GI:27754622
 KEYWORDS
 SOURCE FLI CDNA.
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE 1 (bases 1 to 2162)
 AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shimu, P., Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 TITLE Arabidopsis Full Length cDNA Clones
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 2162)
 AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shimu, P., Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W., Ecker, J.R. and Theologis, A.
 TITLE Direct Submission
 JOURNAL Submitted (15-JAN-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA
 COMMENT RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAFL cDNAs (RAFL cDNA: 'RIKEN Arabidopsis Full-length cDNA'): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PGEC (SSP) Consortium members carried out the sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M., Palm, C.J., Shimu, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W., Ecker, J.R. and Theologis, A.
 Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to this work. Shinzaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC) contributed equally to this work as PIs.

Annotation based on July 2002 version of the Arabidopsis genome submitted to Genbank.

FEATURES
 source 1..2162
 Location/Qualifiers
 1..2162
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /chromosome="5"
 /clone="RAFL14-93-K05 (R20243)"
 /ecotype="Columbia"

```

gene
5'UTR
CDS
/notice="This clone is in a modified pBluescript vector (lambda ps) as a BamHI/XhoI insert."
1..2162
/gene="At5g23570"
1..91
/gene="At5g23570"
92..1969

```

```

/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA022757.1"
/db_xref="Gt:27734623"
/translation="MSSRAAGPMSKKEKNVQGGYRPEVEQLVQGLAGTRLASQDDGGEM
EVIISKNNKPGNTSGKTWVSQNSNPAPAMGCGOQGRGSNTVSGNNVYSGNGNGNGG
IOANSQSGRALSRKYDNNFPVAPPVSPVSPLEGGNGMARGSKOHYAVQFPEVEDD
VDMAEENSDSDALDDSDDLASDDYSDVDSOKSGSKKMKRKKPGSLDLSITFBO
INERPDQNHCRPCQNGPQALIDWYNIHPLIALAPTYGARVYLHEILAEVLEKDIQMKG
ASVLEKIVGQMKLGDEDEKVEYLVPEPMTIMTRITDKDNDKMGAGQELLETF
DKEYLRRARHSYQGGHRMGSLAPESASATYLELRHLRLHRLAMGLDRLAMQGRSM
FSGGQRQVGLFATKQDLIDIFNHSQGGTRLPKLFELKASQEVNVELQISDNDQQLNY
FKNRKSKONAKVLEESLEIMSEKLTAEADNRIVRORTQMOHRENDNEADHDFEYF
MDSIQIHERPRAKEENFEMLOQGRATVYQOOQNIINPSNDQCRAREVVSFTFFP
QEKEMREVEEREMILIKQEKKMEKMKRHHETFDLEKEDFALBQLMYKGLHNE
D"
1970..2162
/genes="At5g23570"
2146
/genes="At5g23570"
/note="not present in genomic sequence"

```

ORIGIN

Query Match	Score	DB	Length
12.1%;	84;	8;	2162;
Best Local Similarity	100.0%;	Prod No	4 38-07.

Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 199 AGTTATTTCTTCCCGAGTCTGACTACTCTCCGCGCTTTAACT 25
|||||
Db 1 AGTTATTTCTTCCCGAGTCTGACTACTCTCCGCGCTTTAACT 60

```
QY      259 TACGTTCTCCGTGGTACTCTGT    282  
        |||||  
Db       61 TAGGTTCTCCGTGGTACTCTGT    84
```

RESULT 5	CNS05TC1	LOCUS	DEFINITION
CNS05TC1	184535 bp	DNA	linear
Homo sapiens chromosome 14 clone R-691G13, WORKING DRAFT SEQUENCE, 8 ordered pieces.			

ACCESSION	AL355100.2
VERSION	GI:9213460
KEYWORDS	HTG; HTGS_PHASE2; HTGS_DRAFT.
SOURCE	Homo sapiens (human)
ORGANISM	Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE	1 (bases 1 to 184535)
AUTHORS	Genoscope.
TITLE	Direct Submission
JOURNAL	Submitted (06-JUN-2000)
COMMENT	On Jul 15, 2000 this sequence

IMPORTANT: This sequence is unfinished and does not necessarily represent the correct sequence. Work on the sequence is in progress and the release of this data is based on the understanding that the sequence may change as work continue. The sequence may be contaminated with foreign sequence from *E. coli*, yeast, vector, phage, etc. . . even if efforts are made to eliminate these contaminating sequences. The following BAC sequence is oriented from the T7 to the SP6 end.

Contigs composition :
1386 bp contig from 1 to 1386
37234 bp contig from 1487 to 38720

1651 bp contig from 38821 to 40471
66410 bp contig from 40572 to 106991
16357 bp contig from 107082 to 123438
2460 bp contig from 123539 to 125998
20897 bp contig from 126099 to 146995
37440 bp contig from 147096 to 184535

```

Overall quality chart :
Range      :      bases
0          :      1377
1 - 9      :      1678
10 - 19    :      1569
20 - 29    :      3627
30 - 39    :      13450
40 - 49    :      13488
50 - 59    :      10047
60 - 69    :      22169
70 - 79    :      54644
80 - 89    :      43202
90 - 99    :      19284
-----
Percentage of bases with a quality value >= 40 : 88 %.
* NOTE: This is a 'working draft' sequence. It currently
* consists of 8 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* been provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.

```

*	1	1366:	config of 1386 bp in length
*	1387	1486:	gap of 100 bp
*	1487	38720:	config of 37224 bp in length
*	38721	38820:	gap of 100 bp
*	38821	40471:	config of 1651 bp in length
*	40472	40571:	gap of 100 bp
*	40572	106961:	config of 66440 bp in length
*	106961	107081:	gap of 100 bp
*	107082	123488:	config of 16357 bp in length
*	123488	123539	gap of 100 bp
*	123539	125998:	config of 2460 bp in length
*	125999	126998:	gap of 100 bp
*	126999	146995:	config of 20897 bp in length
*	146996	147095:	gap of 100 bp
*	147096	184535:	config of 37440 bp in length

```
Location/Qualifiers
1. 184535
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="14"
/clone="R-691G13"
/clone_1fb="RPeI-11"
```

ORIGIN

Query Match Similarity	12.0%	Score	83.4	DB	2	Length	184535
Best Local Similarity	32.3%	Pred. No.	3e-07				
Matches	218	Conservative	0	Mismatches	456	Indels	1
						Gaps	1

Db	QY	2	38591	38650
	ACAAACAAACAAAAA	TTACGACAGTCATGTCGTAGCAATAAATATATAGGGACAAT	61	
	AAAAAAAAAAAAAAAA	AA	38650	

[illegible][illegible]

182 IIGICILICIGAGI AIIICILICLGGAGILCIGACI CACIAICICACI 24

COMMENT

Gaston Cremieux, 91057 Evry cedex, FRANCE
PCR was performed on DNA from transformants of *Arabidopsis thaliana* plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA-derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publications/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.inbio.gen.fr>).

FEATURES

Source

```

1..1407
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/gultivar="wassilewskaja"
/db_xref="taxon:3702"
/clone_id="600D05"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines"
1..1407
/note="T-DNA flanking sequence"
right border"

```

ORIGIN

Query Match	Score	DB	Length
10.5%	72.8	8	1407

[illegible]

RESULT 8	AJ591978	1453 bp	DNA	linear	PLN 23-OCT-2003
LOCUS	AJ591978				
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, right border, clone 59SD01.				
ACCESSION	AJ591978				
VERSION	AJ591978.1	GI:37941602			

KEYWORDS
SOURCE
ORGANIZATION

REFERENCE

AUTHOR

TITLE

103

JOURNAL
OF MEDICINE

PUBME

REFERENCE

AUTHOR
TITLE

JOURNA

COMMENT

KEYWORDS	right border; T-DNA flanking sequence.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana
REFERENCE	1
AUTHORS	Brunaud, V., Balzergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Craud, C., Dekosse, R., Pelletier, G., Lepoint, J., Caboche, M., and Lecharny, A.
TITLE	T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
JOURNAL	EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE	22363535
PUBMED	12446565
REFERENCE	2 (bases 1 to 1453)
AUTHORS	Balzergue, S.
TITLE	Direct Submission
JOURNAL	Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue
PUBMED	Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT	PCR was performed on DNA from transformants of Arabidopsis thaliana plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbsgap.versailles.inra.fr/publiclines/ . This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (http://www.genoplante.com and http://genoplante-info.infobiogen.fr).

FEATURES

Source

```

source
1. 1453
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassilewskaja"
/db_xref="taxon:3702"
/clone="598D01"
/clone_1b="Arabidopsis thaliana T-DNA insertion lines"
misc_feature
1. 1453
/note="T-DNA flanking sequence
right border"

```

ORIGIN

Query Match	10.1%;	Score 70;	DB 8;	Length 1453;
Post Local Similarity	43.4%	Need No	0.0003	

Qy	15	TTTTAGAGCTAAAGAAGCTCCAGTATGTCCTCTCCGCAATTAATTTCTTCCCTCC	216
Db	331	TTTTTNTNTNTT	300
Qy	217	GGAGTCCTGACTCACTACTCTGACTCTCGGCGCTTAACTAGCTTCTCGCTGTTA	276
Db	391	NTT	450
Qy	277	CTGTGAAGTTTTCGCTTGAAGCCTCCGATCGCTCAGCGCATGCAATCTGTGCTCGA	336
Db	451	TTTTNTT	510
Qy	337	TTTCTCTTTTCTCGCGTGAAGAAATTCGCCATAGTCTCGATTGGAAGTTTTGTG	396
Db	511	TT	570
Qy	397	CTATGGGTACTTTTTTCCCTAATTTATTAAGTCTTAGGAAGATACCTGCGCTTAC	456
Db	571	TT	630
Qy	457	TGTTTGTGTCATTTGTGTGCTTTCACCGTTAGTCGATCGAGATTTGAATGT	516
Db	631	TT	690
Qy	517	GAAAAACCTCGTTTTTGGTTTTTGTTGCATATAAATCGAATGATCTACCTTTTGTG	576

[illegible]

RESULT	9
AC138074	
LOCUS	
DEFINITION	66993 bp DNA linear HTG 12-DEC-2002
SEQUENCE	Homo sapiens chromosome 18 clone RP13-907D8 map 18, LOW-PASS SMAPPING.
ACCSSION	AC138074
VERSION	AC138074.1
KEYWORDS	GI:26553381
SOURCE	HTG; HTGS_PHASE0.
ORGANISM	Homo sapiens (human)
	Homo sapiens

```

TITLE      Direct Submission
JOURNAL    Submitted (12-DIC-2002) Whitehead Institute/MIT Center for Genome
COMMENT    Research, 320 Charles Street, Cambridge, MA 02144, USA
           All repeats were identified using RepeatMasker:
           Smt, A.F.A. & Green, P. (1996-1997)
           http://ftp.genome.washington.edu/RM/RepeatMasker.html
           -----
           Genome Center
           Center: Whitehead Institute/ MIT Center for Genome Research
           Center code: MIBR
           Web site: http://www-seg.wi.mit.edu
           Contact: sequence_submissions@genome.wi.mit.edu
           -----
           Project Information
           Center Project name: L28874
           Center Clone name: 907 D 8

```

* NOTE: this record contains 55 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allow
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1	1116:	contig of 1116 bp in length
1117	1216:	gap of 100 bp
1217	2236:	contig of 1120 bp in length
1237	2436:	gap of 100 bp
2437	3561:	contig of 1125 bp in length
3562	3661:	gap of 100 bp
3662	4781:	contig of 1120 bp in length
4782	4881:	gap of 100 bp
4882	6001:	contig of 1120 bp in length
6002	6101:	gap of 100 bp
6102	7244:	contig of 1143 bp in length
7245	7344:	gap of 100 bp
7345	8458:	contig of 1114 bp in length
8459	8558:	gap of 100 bp
8559	9640:	contig of 1082 bp in length
9641	9740:	gap of 100 bp
9741	10886:	contig of 1106 bp in length
10887	10946:	gap of 100 bp
10947	12103:	contig of 1157 bp in length
12104	12203:	gap of 100 bp
12204	13330:	contig of 1117 bp in length
13321	13420:	gap of 100 bp
13421	14520:	contig of 1100 bp in length
14521	14620:	gap of 100 bp
14621	15773:	contig of 1153 bp in length
15774	15873:	gap of 100 bp
15874	16991:	contig of 1118 bp in length
16992	17091:	gap of 100 bp
17092	18340:	contig of 1149 bp in length
18241	18340:	gap of 100 bp
18341	19476:	contig of 1136 bp in length
19477	19576:	gap of 100 bp
19577	20721:	contig of 1145 bp in length
20722	20821:	gap of 100 bp
20822	21942:	contig of 1121 bp in length
21943	22042:	gap of 100 bp
22043	22341:	contig of 1099 bp in length
22342	22341:	gap of 100 bp
22342	24372:	contig of 1131 bp in length
24373	24472:	gap of 100 bp
24473	25627:	contig of 1155 bp in length
25628	25727:	gap of 100 bp
25728	26888:	contig of 1151 bp in length
26889	26988:	gap of 100 bp
26989	28030:	contig of 1042 bp in length
28031	28130:	gap of 100 bp
28131	29303:	contig of 1173 bp in length
29304	29403:	gap of 100 bp
29404	30554:	contig of 1151 bp in length
30555	30654:	gap of 100 bp
30655	31751:	contig of 1097 bp in length
31752	31851:	gap of 100 bp
31852	33966:	contig of 1115 bp in length
33967	33066:	gap of 100 bp
33067	34113:	contig of 1047 bp in length
34114	34213:	gap of 100 bp
34214	35329:	contig of 1116 bp in length
35330	35429:	gap of 100 bp
35430	36574:	contig of 1145 bp in length
36575	36674:	gap of 100 bp
36675	37828:	contig of 1154 bp in length
37829	37928:	gap of 100 bp
37929	39065:	contig of 1135 bp in length
39066	39165:	gap of 100 bp
39166	40288:	contig of 1126 bp in length
40290	40389:	gap of 100 bp
40390	41471:	contig of 1082 bp in length
41472	41571:	gap of 100 bp
41572	42634:	contig of 1063 bp in length
42635	42734:	gap of 100 bp
42735	43904:	contig of 1170 bp in length
43905	44004:	gap of 100 bp
44005	45128:	contig of 1124 bp in length

*	45129	45228: gap of 100 bp	
*	45129	45358: contig of 1130 bp in length	
*	45359	45358: gap of 100 bp	
*	46459	47595: contig of 1137 bp in length	
*	47596	47595: gap of 100 bp	
*	47596	48785: contig of 1090 bp in length	
*	48786	48885: gap of 100 bp	
*	48886	49960: contig of 1075 bp in length	
*	49861	50060: gap of 100 bp	
*	50061	51132: contig of 1072 bp in length	
*	51133	51232: gap of 100 bp	
*	51233	52356: contig of 1124 bp in length	
*	52357	52356: gap of 100 bp	
*	52457	53556: contig of 1100 bp in length	
*	53557	53656: gap of 100 bp	
*	53657	54782: contig of 1126 bp in length	
*	54783	54882: gap of 100 bp	
*	54883	56008: contig of 1126 bp in length	
*	56009	56108: gap of 100 bp	
*	56109	57307: contig of 1199 bp in length	
*	56109	57407: gap of 100 bp	
*	57408	58529: contig of 1122 bp in length	
*	58530	58629: gap of 100 bp	
*	58630	59799: contig of 1170 bp in length	
*	59800	59899: gap of 100 bp	
*	59900	60983: contig of 1084 bp in length	
*	60984	61083: gap of 100 bp	
*	61084	62193: contig of 1110 bp in length	
*	62194	62293: gap of 100 bp	
*	62294	63397: contig of 1104 bp in length	
*	63398	63497: gap of 100 bp	
*	63498	64574: contig of 1077 bp in length	
*	64575	64674: gap of 100 bp	
*	64675	65800: contig of 1126 bp in length	
*	65801	65900: gap of 100 bp	
*	65901	66993: contig of 1093 bp in length	

ORIGIN

Query Match	10.1%;	Score 70;	DB 2;	Length 66993;
Best Local Similarity	41.2%;	Pred. No. 0.00018;		
Matches 214; Conservative	0;	Mismatches 305;	Indels 0;	Gaps 0;

```

/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="18"
/map="18"
/clone="RP13-907D8"
/clone_1fb="RP13 Human Female BAC"

```

QY	13	AAGAGCTAAATAAGC CCGCAATG C C C C C GAGTAAATAT C C C C C	216
Db	56776	TT	56835
QY	217	GGAGTCCTGACTACTACTCTCACTCCGGGCGCTTAAACTAGCTCCCGTGGTTA	276
Db	56836	TTANNNTNNNTTTNTTTNTTTNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	56895
QY	277	CTCTGTAAGTTTCTGCGCTTAGAGCCTCGATCGCTCAGCGATGCAATTCGTGTCGA	336
Db	56896	TTTTTNTNNNNNTTTNTT	56955
QY	337	TTTCTCTTTTCTCGCGAATAATCGCCATATGCTTCGATTTTCGAAGTTTTCGTG	396
Db	56956	TT	57015
QY	397	CTAATGGTACTTTTTTCCCTATATATTATAGTCTTAGAGTAAGATACCTGGCTTAC	456
Db	57016	TT	57075
QY	457	TGTTTTTGTCATTTTGTGTGCTTCACCGTTAGTCGTATCGAGATATTGACTGT	516

[illegible]

RESULT 10	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE
AF592058	LOCUS	1434 bp DNA linear	AF592058	AF592058.1	GI:37941682	right border; T-DNA flanking sequence.	Arabidopsis thaliana (thale cress)	1
		602D05.					Arabidopsis thaliana	
							Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	

TITLE	T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites
JOURNAL	EMBO Rep. 3 (12), 1152-1157 (2002)
MEDLINE	22363535
PUBMED	12446565
REFERENCE	2 (bases 1 to 1434)

AUTHORS	Balzergue, S.
TITLE	Direct Submission
JOURNAL	Submitted (23-OCT-2003) Balzergue S., UMRGV, INRA/CNRS, 2 rue

COMMENT
PCR was performed on DNA from transformants of *Arabidopsis thaliana* plants from INRA (Versailles). The DNA fragment(s) resulting from the PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at <http://dbsgap.versailles.inra.fr/publiclines/>. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (<http://www.genoplante.com> and <http://genoplante-info.inbio.gen.fr>).

FEATURES	Location/Qualifiers
source	1. .1434

```
misc_feature
1..143
/note="T-DNA flanking sequence
right border"
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassilewskaja"
/db_xref="taxon:3702"
/clone="602D05"
/clone_lib="Arabidopsis thaliana T-DNA insertion lines
```

```

ORIGIN
Query Match      10.0%; Score 69.6; DB 8; Length 1434;
Best Local Similarity 43.7%; Pred. No. 0.00036;
Matches 233; Conservative 0; Mismatches 301; Indels 0; Gaps 0

QY      157 TTTTAGAGCTTAATAGCTTCCTCATTTGCTCTCTTCGTCAGTTATTTCTCCCTCC 216
          |||||
Db       503 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 562

```


QY 217 GGAGTCGACTCAGTCTCAGTCTCGGCGCTTAACTTACGTTCCGCTGTTA 276
 Db 563 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 622
 QY 277 CTCTGTAAGTTTCTGCTTAAAGCTCGAATCGGCTACCGCATGATCTGCTGGA 336
 Db 623 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 682
 QY 337 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 396
 Db 683 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 742
 QY 397 CTATGGTACTTTTCTTCCATATTATTAAGTCTTACGTAACGATCGCTTAC 456
 Db 743 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 802
 QY 457 TGTCTTTGTCATTTTGTGTGCTTACCGTTAGTGGCTGATCGAGTATTTGACTGT 516
 Db 803 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 862
 QY 517 GAAAAATCTGCTTTTGTGTTTGTTCATATAAATCGATGATCTTACCTTTGTG 576
 Db 863 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 922
 QY 577 CTGTGATGTTGTTTGTGAGCCTATGCGTGTGCTGTGTTAACTTACGTTCACT 636
 Db 923 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 982
 QY 637 GTGAGTTTGTGATTTGTGATGACGTGCGGTTTCTTGTGCTATAGTGTGT 691
 Db 983 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 1037

RESULT 11
 TSP418778
 LOCUS
 DEFINITION
 TSP418778 3364 bp RNA linear VRL 22-SEP-2003
 Tomato spotted wilt virus NSs gene and N gene, genomic RNA, isolate 10HK96.
 ACCESSION
 AJ418778.1 GI:20429081
 VERSION
 N gene; N protein; NSs gene; NSs protein.
 KEYWORDS
 Tomato spotted wilt virus
 SOURCE
 Tomato spotted wilt virus
 ORGANISM
 Viruses; ssRNA negative-strand viruses; Bunyaviridae; Tospovirus.
 REFERENCE
 1
 Heinze, C., Willingmann, P. and Adam, G.
 Short intergenic regions of the S RNAs of Tomato spotted wilt
 tospovirus - not a species characteristic
 Unpublished
 2 (bases 1 to 3364)
 JOURNAL
 Adam, G.
 Direct Submission
 Submitted (05-NOV-2001) Adam G., Pflanzenschutz, Institut fuer
 Angewandte Botanik, Ohnhorststrasse 18, 22609 Hamburg, GERMANY
 FEATURES
 SOURCE
 1..3364
 /organism="Tomato spotted wilt virus"
 /viralion
 /mol_type="genomic RNA"
 /isolate="10HK96"
 /specific_host="Virginia tobacco"
 /db_xref="taxon:11613"
 /country="Bulgaria"
 89..1492
 /gene="NSs"
 89..1492
 /gene="NSs"
 /codon_start=1
 /product="NSs protein"
 /protein_id="CAD11445.1"
 /db_xref="GI:20429082"
 /db_xref="SPTREMBL:O8JXK4"

ORIGIN

/translation="MSSSVYESIIQTRASVWSGTASGAKVDSYWIHELGTSLVOT
 QLYSDSKSSFGYTKAVGNLPCEESEIIISQHYIPIFDIDPSIINDISVLAISCS
 MNTVNGVKGHGLTKYLSPOHLSIGSTNMRSIDIKRFOLOEEDIIPNRYIAAANG
 SLSCVKEHTYKITECTCNQALGKRVNYSPPNNVSEIYSPFKNSNOVSENNRYSLAV
 RSLTSLSENNIMENSOAFVASTDSHPKSLMRYRVKVARQVSIQCLFVAGDETKT
 FYLSIVCIENHNSVEFALNLSVICIKQPLRKARSELSWSDLEKEPNIVHDSY
 PORIYVALLTEHTSFAOVLCSNLOEDEVITYLLNNYELTPKDLGERTVLSYSDICR
 KYFLSKTLECLPSTOTFMSYLDISIQIPSWKIDFARGGIKISIPQVSAVSKLKLDSG
 IKKRSKISDVSAGSKS"
 complement (2436..3212)
 /gene="N"
 complement (2436..3212)
 /gene="N"
 /function="structural protein"
 /codon_start=1
 /product="N protein"
 /protein_id="CAD11446.1"
 /db_xref="GI:20429083"
 /db_xref="GOA:O8JXK3"
 /db_xref="SPTREMBL:O8JXK3"
 /translation="MSKVKLTRESIYVALLTQKOLFEEDQNLVAENFKTCLEMLDQ
 IKRNSVISCLEFPKRNOSIMKVKOSDFTGKITIKTSRDIATDWTFRLLDSMIRY
 RLVEIGNSENLNTIKSKISHPILQAYGLPDKASVRLAIMGSLPLIASVDSPE
 MISVVALIYQDANKDGLIDPKYDITGALGKVCVTLSKSAFENAEQVKKGEYAL
 LSSNPNKSGIAMGHVSETLNKFEVMEFVKOKKLTALA"

Query Match 10.0%; Score 69.2; DB 14; Length 3364;
 Best Local Similarity 45.7%; Pred. No. 0.00039;
 Matches 281; Conservative 0; Mismatches 330; Indels 4; Gaps 1;

QY 39 AATAAATTAAATAGTGGAAACAATTAAAGTAAAGCAAAAAAGAAAAAGCTACAAAA 98
 Db 1629 AATCAAAAATGAAACAAAATCAAAAATGAAACAAAATGAAACAAAATGAAACAAA 1688
 QY 99 TGAATAACAAATCAAACTGAATGAAATTTGAGATCGAATTCGAAATAAGAGCGCTT 158
 Db 1689 CAAAAAAATTAATAAAACAAAAAACAAAAAGATCCGAAAGGCAATTTGGCCAAAT 1748
 QY 159 TTAGAGCTTAATTAAGCTTCATTTGTCCTTCTCTGCTGATTAATTTCTCTCCGG 218
 Db 1749 TTGGGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 1808
 QY 219 AGTCTGACTCACTACTGACTCTCGCGCGCTTAACTTAAGTTCCTCCGCTTACT 278
 Db 1809 TTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTAT 1868
 QY 279 CTGTAGTTTCTGCTTAAAGCTCGGATCGCTCACCGATGATCTGCTCGATT 338
 Db 1869 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1924
 QY 339 TCTCTTTTCTCGCTGAAAAATGCGCTAATGCTTCGATTTGAAAGTTTGTGCT 398
 Db 1925 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 1984
 QY 399 ATGGGTACTTTTCTCCATAATTATTAAGTCTTAAGTAAAGATACCGCTCTTACG 458
 Db 1985 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2044
 QY 459 TTTTCTTCAATTTGTGTCCTTACCGTTAGTGGCTGATCGAGTATTTGACTGGA 518
 Db 2045 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2104
 QY 519 AAAATCCTGCTTTTGTGTTTGTGTTGATATAAATCGAATGATCTTGTGCT 578
 Db 2105 ATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2164
 QY 579 TTGATGTTGTTTGTGAGCTTAAAGCTGTGCTGTGTTATTAACCTGATCGATGCT 638
 Db 2165 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 2224
 QY 639 GAATTTGAGATTTT 653

DB 2225 GTTATTTTGTGTT 2239

RESULT 12
AC022680
LOCUS
DEFINITION
AC022680 Homo sapiens chromosome 8 clone RP11-511E14 map 8, LOW-PASS
SEQUENCE SAMPLING.
AC022680
AC022680.2 GI:9143927
KEYWORDS
HTG, HTGS_PHASE0.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1. (bases 1 to 88932)
Birren, B., Linton, L., Nusbaum, C. and Landier, E.
Homo sapiens chromosome 8, clone RP11-511E14
Unpublished
2. (bases 1 to 88932)
Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
Anderson, S., Baldwin, J., Barna, N., Beckerly, R., Beda, F.,
Boguslavsky, L., Bouckgalter, B., Brown, A., Burkett, G., Castle, A.,
Choepel, Y., Colangelo, M., Collins, S., Collymore, A., Cooke, P.,
Dearellano, K., Dewar, K., Domino, M., Doyle, M., Feneator, J.,
Ferrera, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
Gardina, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J. C., Johnson, R., Jones, C., Kann, L., Karatae, A., Klein, J.,
Landers, T., Lehocsky, J., Levine, R., Liu, C., Liu, G., Locke, K.,
Macdonald, P., Margulis, N., McEwan, P., McGurk, A., McKernan, K.,
Mcpheters, R., Meldrum, J., Meneus, L., Morrow, J., Naylor, J.,
Norman, C. H., O'Connor, T., O'Donnell, P., Oliver, T. M., Peterson, K.,
Pierre, N., Pisan, C., Pollara, V., Raymond, C., Riley, R., Rothman, D.,
Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thumann, N.,
Stojanovic, N., Sudramanian, A., Talamas, J., Testaye, S., Theodore, J.,
Tirrell, A., Vassiliev, H., Viel, R., Vo, A., Wu, X., Wyman, D., Ye, W. J.,
Zimmer, A. and Zody, M.
Direct Submission
Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Jul 13, 2000 this sequence version replaced g1:6910620.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5106
Center clone name: 511_E_14

* NOTE: This record contains 86 individual
* sequencing reads that have not been assembled into
* contigs. Runs of N are used to separate the reads
* and the order in which they appear is completely
* arbitrary. Low-pass sequence sampling is useful for
* identifying clones that may be gene-rich and allows
* overlap relationships among clones to be deduced.
* However, it should not be assumed that this clone
* will be sequenced to completion. In the event that
* the record is updated, the accession number will
* be preserved.

1
934 1033: contig of 933 bp in length
* 934 1033: gap of 100 bp
* 1034 1961: contig of 928 bp in length
* 1962 2061: gap of 100 bp
* 2062 3005: contig of 944 bp in length
* 3006 3105: gap of 100 bp
* 3106 4066: contig of 961 bp in length
* 4067 4166: gap of 100 bp
* 4167 5120: contig of 954 bp in length

5121 5220: gap of 100 bp
* 5221 6142: contig of 922 bp in length
* 6143 6242: gap of 100 bp
* 6243 7207: contig of 965 bp in length
* 7208 7307: gap of 100 bp
* 7308 8234: contig of 927 bp in length
* 8235 8334: gap of 100 bp
* 8335 9276: contig of 942 bp in length
* 9277 9376: gap of 100 bp
* 9377 10329: contig of 953 bp in length
* 10330 10429: gap of 100 bp
* 10430 11370: contig of 941 bp in length
* 11371 11470: gap of 100 bp
* 11471 12439: contig of 969 bp in length
* 12440 12539: gap of 100 bp
* 12540 13474: contig of 935 bp in length
* 13475 13574: gap of 100 bp
* 13575 14538: contig of 964 bp in length
* 14539 14638: gap of 100 bp
* 14639 15595: contig of 957 bp in length
* 15596 15695: gap of 100 bp
* 15696 16623: contig of 928 bp in length
* 16624 16723: gap of 100 bp
* 16724 17659: contig of 936 bp in length
* 17660 17759: gap of 100 bp
* 17760 18702: contig of 943 bp in length
* 18703 18802: gap of 100 bp
* 18803 19757: contig of 955 bp in length
* 19758 19857: gap of 100 bp
* 19858 20787: contig of 930 bp in length
* 20788 20887: gap of 100 bp
* 20888 21836: contig of 949 bp in length
* 21837 21936: gap of 100 bp
* 21937 22865: contig of 925 bp in length
* 22866 22965: gap of 100 bp
* 22966 23906: contig of 941 bp in length
* 23907 24006: gap of 100 bp
* 24007 24929: contig of 923 bp in length
* 24930 25029: gap of 100 bp
* 25030 25953: contig of 924 bp in length
* 25954 26053: gap of 100 bp
* 26054 26992: contig of 939 bp in length
* 26993 27092: gap of 100 bp
* 27093 28049: contig of 957 bp in length
* 28050 28149: gap of 100 bp
* 28150 29108: contig of 959 bp in length
* 29109 29208: gap of 100 bp
* 29209 30124: contig of 916 bp in length
* 30125 30224: gap of 100 bp
* 30225 31124: contig of 900 bp in length
* 31125 31224: gap of 100 bp
* 31225 32179: contig of 955 bp in length
* 32179 32279: gap of 100 bp
* 32279 33207: contig of 928 bp in length
* 33208 33307: gap of 100 bp
* 33308 34249: contig of 942 bp in length
* 34250 34349: gap of 100 bp
* 34350 35289: contig of 940 bp in length
* 35290 35389: gap of 100 bp
* 35390 36327: contig of 938 bp in length
* 36328 36427: gap of 100 bp
* 36428 37353: contig of 926 bp in length
* 37354 37453: gap of 100 bp
* 37454 38408: contig of 955 bp in length
* 38409 38508: gap of 100 bp
* 38509 39427: contig of 919 bp in length
* 39428 39527: gap of 100 bp
* 39528 40459: contig of 932 bp in length
* 40460 40559: gap of 100 bp
* 40560 41483: contig of 924 bp in length
* 41484 41583: gap of 100 bp
* 41584 42543: contig of 960 bp in length
* 42544 42643: gap of 100 bp

```

* 42644 43582: contig of 939 bp in length
* 43583 43682: gap of 100 bp
* 44639 44639: contig of 957 bp in length
* 44640 44739: gap of 100 bp
* 44740 45670: contig of 931 bp in length
* 45671 45770: gap of 100 bp
* 45771 46715: contig of 945 bp in length
* 46716 46815: gap of 100 bp
* 46816 47746: contig of 931 bp in length
* 47747 47846: gap of 100 bp
* 47847 48772: contig of 926 bp in length
* 48773 48872: gap of 100 bp
* 48873 49804: contig of 932 bp in length
* 49805 49904: gap of 100 bp
* 49905 50849: contig of 945 bp in length
* 50850 50949: gap of 100 bp
* 50950 51842: contig of 893 bp in length
* 51843 51942: gap of 100 bp
* 51943 52882: contig of 940 bp in length
* 52883 52982: gap of 100 bp
* 52983 53928: contig of 946 bp in length
* 53929 54028: gap of 100 bp
* 54029 54988: contig of 960 bp in length
* 54989 55088: gap of 100 bp
* 55089 56034: contig of 946 bp in length
* 56035 56134: gap of 100 bp
* 56135 57063: contig of 929 bp in length
* 57064 57163: gap of 100 bp
* 57164 58097: contig of 934 bp in length
* 58098 58197: gap of 100 bp
* 58198 59135: contig of 938 bp in length
* 59136 59235: gap of 100 bp
* 59236 60161: contig of 926 bp in length
* 60162 60261: gap of 100 bp
* 60262 61180: contig of 919 bp in length
* 61181 61280: gap of 100 bp
* 61281 62213: contig of 933 bp in length
* 62214 62313: gap of 100 bp
* 62314 63241: contig of 928 bp in length
* 63242 63341: gap of 100 bp
* 63342 64242: contig of 901 bp in length
* 64243 64342: gap of 100 bp
* 64343 65301: contig of 959 bp in length
* 65302 65401: gap of 100 bp
* 65402 66365: contig of 964 bp in length
* 66366 66465: gap of 100 bp
* 66466 67381: contig of 916 bp in length
* 67382 67481: gap of 100 bp
* 67482 68455: contig of 974 bp in length
* 68456 68555: gap of 100 bp
* 68556 69479: contig of 924 bp in length
* 69480 69580: gap of 100 bp
* 69581 70512: contig of 933 bp in length
* 70513 70612: gap of 100 bp
* 70613 71530: contig of 918 bp in length
* 71531 71630: gap of 100 bp
* 71631 72664: contig of 934 bp in length
* 72665 72864: gap of 100 bp
* 72865 73601: contig of 937 bp in length

```

Query Match
Best Local Similarity 45.0%; Pred. No. 0.00027; Length 88932;
Matches 165; Conservative 0; Mismatches 202; Indels 0; Gaps 0;

```

QY 325 TTCTGTGCTGATTTCTTTTCTGCTGGAATAATGCCCTATGTTCTGATTCG 384
DB 68986 TTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 69045
QY 385 AAGCTTTTGTGCTATGAGTTACTTTTTCCTATATTTATATAGTTCTTACGTA 444
DB 69046 TTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 69105
QY 445 CCTGGCTTACTGTTTGTGTCATTTGTGTGCTTACCGCTTACGCTATGCGGA 504

```

```

DB 69106 TTTNTTTNNNTTTTNGNGNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 69165
QY 505 GATATGACTGTGAAAATCCTTCGTTTGTGTTTGTTCATATAATGATGAT 564
DB 69166 TNGTTTGTGTTTGTGTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 69225
QY 565 CTACCTTTTGTGCTTATGATGTTGTTTGTGAGCCTATGCGTTGTGCTGTAA 624
DB 69226 TTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 69285
QY 625 TCACGTTGATGCTGATTTGAGATTTGAGATGAGACCTGCGGTTCTTGTGCG 684
DB 69286 TTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 69345
QY 685 AAGTTGT 691
DB 69346 TTTTNTT 69352

```

```

RESULT 13
AX346599
LOCUS AX346599 6668 bp DNA linear PAT 01-FEB-2002
DEFINITION Sequence 1670 from Patent WO0200928.
ACCESSION AX346599
VERSION AX346599.1 GI:18494485
KEYWORDS
SOURCE
ORGANISM
synthetic construct
synthetic construct
artificial sequences.
REFERENCE
1
AUTHORS Olek, A.; Piepenbrock, C. and Berlin, K.
TITLE Diagnosis of diseases associated with the immune system
JOURNAL Patent: WO 0200928-A 1670 03-JAN-2002;
Epigenomics AG (DE)
FEATURES
location/Qualifiers
source
1..6668
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="chemically treated genomic DNA (Homo sapiens)"
ORIGIN

```

Query Match
Best Local Similarity 45.8%; Pred. No. 0.00043; Length 6668;
Matches 238; Conservative 0; Mismatches 282; Indels 0; Gaps 0;

```

QY 156 GTTTAAGCTTAATAAGCTTCTCATTTGCTCTCTGTCAGTTATTTCTCTC 215
DB 2742 GTTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTT 2801
QY 216 CGAGTCTGACCTACACCTCTGCTCGGGGCTTAACTTACCTTCCGCTGT 275
DB 2802 TTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 2861
QY 276 ACTCTGATGATTTCTGCTTAAAGCTCCGATCCGCTACCGCATTCGTGCTG 335
DB 2862 TTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 2921
QY 336 ATTTCCTTTTCTGCTGGAATAATGCCCTAATGTTCTGATTTGGAAGTTTGT 395
DB 2922 TTTTGTGTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 2981
QY 396 GCTATGCTTACTTTTCCCTATTTTATATATATGTTCTTATGATACATACCT 455
DB 2982 TTTTGTGTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 3041
QY 456 CTGTTTGTGCTATTTGTTGCTTCAACCGTTAGTGCCTGATGAGATTTGACG 515
DB 3042 TTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTTNTTTT 3101
QY 516 TGAATAATCCTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 575

```

```
Db      3102 TGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGT 3161
Oy      576 GCTTGAATGTTGTTTGTAGCGTATGCGTGTGCTGTATACCTACGTCATG 635
Db      3162 TTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3221
Oy      636 TGTGATTTTGAGATTTTGCTAGTGTGCTGTTCTTT 675
Db      3222 TTTTGTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 3261

RESULT 14
AC022851      81120 bp   DNA      linear   HTG 13-JUL-2000
LOCUS      AC022851
DEFINITION Homo sapiens chromosome 11 clone RP11-284N16 map 11, LOW-PASS
SEQUENCE SAMPLING.
AC022851
AC022851.2 GI:9158172
VERSION      AC022851.2
KEYWORDS     HTG; HTGS_PHASE0.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE    1. (bases 1 to 81120)
AUTHORS      Birren,B., Linton,L., Nuebaum,C. and Lander,E.
TITLE        Homo sapiens chromosome 11, clone RP11-284N16
JOURNAL      Unpublished
REFERENCE    2. (bases 1 to 81120)
AUTHORS      Birren,B., Linton,L., Nuebaum,C., Lander,E., Abraham,H., Allen,N.,
              Anderson,S., Baldwin,J., Barua,N., Beckerly,R., Bede,F.,
              Boguslavsky,L., Boukhalter,B., Brown,A., Burkett,G., Castle,A.,
              Choquel,Y., Colangelo,M., Collins,S., Collymore,A., Cooke,P.,
              Dearlano,K., Dewar,K., Domini,M., Doyle,M., Fenesfor,J.,
              Ferreira,P., Fitzhugh,W., Forrest,C., Gage,D., Galagan,J.,
              Gardyna,S., Grant,G., Hagos,B., Heatord,A., Horton,L.,
              Howland,J.C., Johnson,R., Jones,C., Kann,L., Karatas,A., Klein,J.,
              Landers,T., Iehoczky,J., Levine,R., Lieu,C., Liu,G., Locke,K.,
              Macdonald,P., Marguis,N., McEwan,P., McGurk,A., McKernan,K.,
              McPheeters,R., Meldrum,J., Meneus,L., Morrow,J., Naylor,J.,
              Norman,C.H., O'Connor,T., O'Donnell,P., Olivar,T.M., Peterson,K.,
              Plerre,N., Pisanl,C., Pollara,V., Raymond,C., Riley,R., Rothman,D.,
              Roy,A., Santos,R., Severy,P., Spencer,B., Stange-Thoman,N.,
              Stojanovic,N., Subramanian,A., Talamas,J., Tesfaye,S., Theodore,J.,
              Tirrell,A., Vassiliev,H., Viel,R., Vo,A., Wu,X., Wyman,D., Ye,W.J.,
              Zimmer,A. and Zody,M.
              Direct Submission
              Submitted (06-FEB-2000) Whitehead Institute/MIT Center for Genome
              Research, 320 Charles Street, Cambridge, MA 02141, USA
              On Jul 13, 2000 this sequence version replaced gi:6922196.
              All repeats were identified using RepeatMasker:
              Smit, A.P.A. & Green, P. (1996-1997)
              http://ftp.genome.washington.edu/RM/RepeatMasker.html
              ----- Genome Center
              Center: Whitehead Institute/ MIT Center for Genome Research
              Center code: WIBR
              Web site: http://www-seq.wi.mit.edu
              Contact: sequence_submissions@genome.wi.mit.edu
              ----- Project Information
              Center project name: L6170
              Center clone name: 284_N_16
              -----
              * NOTE: This record contains 82 individual
              * sequencing reads that have not been assembled into
              * contigs. Runs of N are used to separate the reads
              * and the order in which they appear is completely
              * arbitrary. Low-pass sequence sampling is useful for
              * identifying clones that may be gene-rich and allows
              * overlap relationships among clones to be deduced.
              * However, it should not be assumed that this clone
              * will be sequenced to completion. In the event that
              * the record is updated, the accession number will
              * be preserved.
              1          910: contig of 910 bp in length
              911          1010: gap of 100 bp
              1011          1910: contig of 900 bp in length
              1911          2010: gap of 100 bp
              2011          2917: contig of 907 bp in length
              2918          3017: gap of 100 bp
              3018          3921: contig of 904 bp in length
              3922          4021: gap of 100 bp
              4022          4903: contig of 882 bp in length
              4904          5003: gap of 100 bp
              5004          5874: contig of 871 bp in length
              5875          5974: gap of 100 bp
              5976          6860: contig of 886 bp in length
              6861          6960: gap of 100 bp
              6961          7856: contig of 896 bp in length
              7857          7956: gap of 100 bp
              7957          8860: contig of 904 bp in length
              8861          8960: gap of 100 bp
              8961          9822: contig of 862 bp in length
              9823          9922: gap of 100 bp
              9923          10818: contig of 896 bp in length
              10819          10918: gap of 100 bp
              10919          11817: contig of 899 bp in length
              11818          11917: gap of 100 bp
              11917          12797: contig of 880 bp in length
              12798          12897: gap of 100 bp
              12898          13776: contig of 879 bp in length
              13777          13876: gap of 100 bp
              13877          14756: contig of 880 bp in length
              14757          14856: gap of 100 bp
              14857          15788: contig of 882 bp in length
              15789          15838: gap of 100 bp
              15839          16723: contig of 885 bp in length
              16724          16823: gap of 100 bp
              16824          17689: contig of 876 bp in length
              17689          17799: gap of 100 bp
              17799          18699: contig of 900 bp in length
              18700          18799: gap of 100 bp
              18799          19689: contig of 890 bp in length
              19690          19789: gap of 100 bp
              19789          20687: contig of 908 bp in length
              20688          20797: gap of 100 bp
              20798          21638: contig of 841 bp in length
              21639          21738: gap of 100 bp
              21739          22633: contig of 895 bp in length
              22634          22733: gap of 100 bp
              22734          23621: contig of 888 bp in length
              23622          23721: gap of 100 bp
              23722          24617: contig of 896 bp in length
              24618          24717: gap of 100 bp
              24718          25612: contig of 895 bp in length
              25613          25712: gap of 100 bp
              25713          26579: contig of 867 bp in length
              26580          26679: gap of 100 bp
              26680          27577: contig of 898 bp in length
              27578          27677: gap of 100 bp
              27678          28569: contig of 892 bp in length
              28570          28659: gap of 100 bp
              28670          29681: contig of 912 bp in length
              29682          29681: gap of 100 bp
              29682          30578: contig of 897 bp in length
              30579          30678: gap of 100 bp
              30679          31580: contig of 902 bp in length
              31581          31680: gap of 100 bp
              31681          32568: contig of 888 bp in length
              32569          32668: gap of 100 bp
              32669          33562: contig of 894 bp in length
              33563          33662: gap of 100 bp
              33663          34573: contig of 911 bp in length
              34574          34673: gap of 100 bp
              34674          35579: contig of 906 bp in length
              35579          35679: gap of 100 bp
              35680          36547: contig of 868 bp in length
              36548          36647: gap of 100 bp
```

36648	37646:	cont: of 899	bp	in	length
37547	37646:	gap of 100	bp		
37647	38553:	cont: of 907	bp	in	length
38554	38653:	gap of 100	bp		
38654	39526:	cont: of 873	bp	in	length
39527	39626:	gap of 100	bp		
39627	40521:	cont: of 895	bp	in	length
40522	40621:	gap of 100	bp		
40622	41509:	cont: of 888	bp	in	length
41510	41609:	gap of 100	bp		
41610	42520:	cont: of 911	bp	in	length
42521	42620:	gap of 100	bp		
42621	43507:	cont: of 887	bp	in	length
43508	43607:	gap of 100	bp		
43608	44234:	cont: of 627	bp	in	length
44235	44334:	gap of 100	bp		
44335	45212:	cont: of 878	bp	in	length
45213	45312:	gap of 100	bp		
45313	46197:	cont: of 885	bp	in	length
46198	46297:	gap of 100	bp		
46298	47174:	cont: of 877	bp	in	length
47175	47274:	gap of 100	bp		
47275	48171:	cont: of 837	bp	in	length
48172	48271:	gap of 100	bp		
48272	49157:	cont: of 886	bp	in	length
49158	49257:	gap of 100	bp		
49258	50252:	cont: of 895	bp	in	length
50153	50252:	gap of 100	bp		
50253	51155:	cont: of 903	bp	in	length
51156	51255:	gap of 100	bp		
51256	52145:	cont: of 890	bp	in	length
52146	52245:	gap of 100	bp		
52246	53156:	cont: of 911	bp	in	length
53157	53255:	gap of 100	bp		
53257	54139:	cont: of 883	bp	in	length
54140	54239:	gap of 100	bp		
54240	55149:	cont: of 910	bp	in	length
55150	55249:	gap of 100	bp		
55250	56130:	cont: of 881	bp	in	length
56131	56230:	gap of 100	bp		
56231	57112:	cont: of 882	bp	in	length
57113	57212:	gap of 100	bp		
57213	58095:	cont: of 883	bp	in	length
58096	58195:	gap of 100	bp		
58196	59098:	cont: of 903	bp	in	length
59099	59198:	gap of 100	bp		
59199	60093:	cont: of 895	bp	in	length
60094	60193:	gap of 100	bp		
60194	61096:	cont: of 903	bp	in	length
61097	61196:	gap of 100	bp		
61197	62069:	cont: of 873	bp	in	length
62070	62169:	gap of 100	bp		
62170	63036:	cont: of 867	bp	in	length
63037	63136:	gap of 100	bp		
63137	64048:	cont: of 912	bp	in	length
64049	64148:	gap of 100	bp		
64149	65036:	cont: of 888	bp	in	length
65037	65136:	gap of 100	bp		
65137	66020:	cont: of 884	bp	in	length
66021	66120:	gap of 100	bp		
66121	66984:	cont: of 864	bp	in	length
66985	67084:	gap of 100	bp		
67085	68014:	cont: of 930	bp	in	length
68015	68114:	gap of 100	bp		
68115	69032:	cont: of 918	bp	in	length
69033	69132:	gap of 100	bp		
69133	70016:	cont: of 886	bp	in	length

Query Match	9.9%	Score 68.8;	DB 2;	Length 81120;
Best Local Similarity	41.8%;	Pred. No. 0.0003;		
Matches 211; Conservative	0;	Mismatches 294;	Indels 0;	Gaps 0

175 TTCCCATTTTCCTTCTTCGTAGTTATTTTCTTCCTCCGAGTCCTGACTCATAC 234

Db	9261	TTTTTTTNNNNNTTTTTTTTTNNNNNTTTTTTTTTTTTTNNNTTTTTTTTTTTTT	9320
OY	235	TCTCAGTCGCGGGCTTTAAACTTAGCTCTCCGTGTTACCTCTGAAGTTTTCGCC	294
Db	9321	TTNNT	9380
OY	295	TTAAGCCTCGAATGCCTCACGCAATCTGTCGATTCTCTTTTCTTGCT	354
Db	9381	TTTTTTTTTTTTTTTTTTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	9440
OY	355	GAAAAATGCCCCAATGTTCTCGATTCGAAGTTTGTGCTAGGTTACTTTTTTC	414
Db	9441	TTT	9500
OY	415	CCATATTTAATAGTCTTAGAGTACATACCTGCGCTACTGTTTTGTCAATTTGT	474
Db	9501	TTTTTTTTNNNNNTTTTTTTTTNNNTTTTTTTTTTTTTTTTTTTTTNNNTTTTTT	9560
OY	475	TGNCCTTCAACGGTTTAGTCGCTGATCGAGATTTAGTGTAAAATCCTTCGTTTT	534
Db	9561	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTNNNTNNNNNNNTTTTTTTT	9620
OY	535	TGTTTTTGTTCATAATAATCGAATGATCTACCTTTGTGCTTGATGTTGTTTTT	594
Db	9621	TTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTTNNNTTTTTTTTTTGTTTTTTTTTT	9680
OY	595	GAGCCTATGCGTTGTGGCTTGTTAATACTCACGTCAGTGTGAGATTTGAGAATTTTG	654
Db	9681	TTTTTTTTNNNTGTTTTTTTTTNGTTTTTGTGTTGNNNTTGTNNNTGTGGGTGTTG	9740
OY	655	GTATGACTGTGGGTTCTTTGTGTG 679	
Db	9741	TGTGGGGGCTGTGGGTGGGG 9765	

	RESULT	15
AJ592180	LOCUS	
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence, right border, clone 60dH03.	810 bp DNA linear PLN 23-OCT-2003
ACCESSION	AJ592180	
VERSION	AJ592180.1 GI:37941804	
KEYWORDS	right border; T-DNA flanking sequence.	
SOURCE	Arabidopsis thaliana (thale cress)	
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophytes; Magnoliophyta; eudicotyledons; core eudicots; rosids, euroside II; Brassicales; Brassicaceae; Arabidopsi-	
REFERENCE	1 Brunaud, V., Balergue, S., Dubreucq, B., Aubourg, S., Samson, F., Chauvin, S., Bechtold, N., Crund, C., Dekosse, R., Pelletier, G., Lepoint, L., Caboche, M. and Leclarny, A. T-DNA integration into the Arabidopsis genome depends on sequences of pre-insertion sites EMBO Rep. 3 (12), 1152-1157 (2002)	
TITLE	JOURNAL MEDLINE PUBLISHED IN	
PUBMED	22363535	
REFERENCE	2 (bases 1 to 810) Balergue, S. Direct Submmission	
AUTHORS	Gaston Cremieux, 91057 Evry cedex, FRANCE	
TITLE	InRA Versailles) The DNA fragment(s) resulting from plants from InRA (Versailles). The PCR were directly sequenced from the left or the right border to determine the genomic sequence flanking the insertion. T-DNA derived sequences were removed. Information to order the corresponding mutant line and a link to a database providing a graphical display of the insertion site are available at http://dbseq.versailles.inra.fr/publicines/. This sequence has been generated in the framework of the French plant genomics program 'Genoplante' (http://www.genoplante.com) and	
JOURNAL	BALZERGUE, S., UMRGV, INRA/CNRS, 2 rue	

FEATURES http://genoplante-info.infobiogen.fr).

source Location/Qualifiers

1..810
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Wassilewskija"
/db_xref="taxon:3702"
/clone="606H03"

misc_feature 1..810
/note="T-DNA flanking sequence"

right border"

ORIGIN

Query Match 9.8%; Score 68.4; DB 8; Length 810;

Best Local Similarity 41.2%; Pred. No. 0.00069;

Matches 213; Conservative 0; Mismatches 304; Indels 0; Gaps 0;

```
QY 175 TTCCGATTTGCTCTCTCTGTCAGTTTATTTCTTCCTCCGAGTCTGACTCATC 234
   |||||
Db 142 TTTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 201
   |||||
QY 235 TCTCACTCTCCGGCGCTTAACCTACGTCCTCCGCTTACTCTGAAGTTTCTGCC 294
   |||||
Db 202 TTTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 261
   |||||
QY 295 TTAGAGCCCTCCGATCCGCTACCGCATGCTGCTGCTGATTTCTTTTCTTGGCT 354
   |||||
Db 262 TTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 321
   |||||
QY 355 GGAATAATGCCCCTAATGCTCGATTTGCAAGTTTGTGCTATGAGTTACTTTTTC 414
   |||||
Db 322 TTTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 381
   |||||
QY 415 CCTAATTTTAACTCTTAGTAAAGAACTGCGCTTACTGTTTGTCAATTTGT 474
   |||||
Db 382 NTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 441
   |||||
QY 475 TGTGCTTTGACGCTTAGTGTGCTGATCCGAGTATTGACGTGAAATGCTGTTTT 534
   |||||
Db 442 TTTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 501
   |||||
QY 535 TGGTTTTTGTTCATATAATCGAATGATCTACCTTTGCTTGATGTTGTTTTT 594
   |||||
Db 502 TTTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 561
   |||||
QY 595 GAGCCTTAGCGTGTGGCTGTAACTTCACTTCAATGATGATTTGAGATTTTG 654
   |||||
Db 562 TTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTT 621
   |||||
QY 655 GTAGTGACTGTGGGTTCTTTGCTGCTAAGGTGT 691
   |||||
Db 622 TTTTCTTTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNTTNT 658
   |||||
```

Search completed: April 8, 2004, 13:33:39
Job time : 1818.04 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 19:34:49 ; Search time 165.206 Seconds
(without alignments)
11001.196 Million cell updates/sec

Title: US-10-030-829-1
Perfect score: 3275
Sequence: 1 gcaacaacaacaataa.....caaatataggtgacacat 3275

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: Issued Patents NA:*
2: /cgn2_6/prodata/2/ina/5A_COMB.seq:*
3: /cgn2_6/prodata/2/ina/5B_COMB.seq:*
4: /cgn2_6/prodata/2/ina/6A_COMB.seq:*
5: /cgn2_6/prodata/2/ina/6B_COMB.seq:*
6: /cgn2_6/prodata/2/ina/6C_COMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	85.6	2.6	7218	1	US-08-232-463-14	Sequence 14, Appl
2	74.4	2.3	272	4	US-09-313-294A-121	Sequence 121, App
3	72.2	2.2	274	4	US-09-313-294A-463	Sequence 463, App
4	68.6	2.1	7218	1	US-08-232-463-14	Sequence 14, Appl
5	52.6	1.6	7286	3	US-09-331-581-3	Sequence 3, Appl
6	52.6	1.6	7938	3	US-09-331-581-14	Sequence 14, Appl
7	50.6	1.5	396	4	US-09-640-173-53	Sequence 53, Appl
8	50.6	1.5	396	4	US-09-713-550-53	Sequence 53, Appl
9	50.6	1.5	10619	4	US-10-204-708-4	Sequence 4, Appl
10	50.2	1.5	5562	4	US-10-204-708-63	Sequence 63, Appl
11	49.4	1.5	19124	2	US-08-487-826B-13	Sequence 13, Appl
12	49.2	1.5	6040	4	US-10-204-708-69	Sequence 69, Appl
13	48.6	1.5	51259	3	US-08-781-891-209	Sequence 209, App
14	48.6	1.5	51259	4	US-09-618-166-209	Sequence 14, App
15	47.8	1.5	2447	2	US-09-014-969-14	Sequence 14, App
16	47.4	1.4	5666	4	US-10-204-708-29	Sequence 29, Appl
17	47.4	1.4	8607	4	US-10-204-708-72	Sequence 72, Appl
18	47.4	1.4	19233	4	US-10-204-708-45	Sequence 45, Appl
19	47.2	1.4	8961	4	US-10-204-708-79	Sequence 79, Appl
20	47.2	1.4	6617	4	US-09-976-594-268	Sequence 268, App
21	46	1.4	3489	2	US-08-728-323A-1	Sequence 1, Appl
22	46	1.4	3489	4	US-09-298-568-1	Sequence 1, Appl
23	46	1.4	3489	4	US-09-410-399-1	Sequence 1, Appl
24	46	1.4	32207	2	US-08-770-379-20	Sequence 20, Appl
25	46	1.4	32207	2	US-08-757-669A-20	Sequence 20, Appl
26	46	1.4	32207	4	US-09-230-371A-20	Sequence 20, Appl
27	45.6	1.4	289	3	US-09-007-005-17	Sequence 17, Appl

C 28	45.6	1.4	289	3	US-09-244-796-17	Sequence 17, Appl
C 29	45.4	1.4	740	3	US-08-998-416-563	Sequence 563, Appl
C 30	45.4	1.4	11049	4	US-10-204-708-23	Sequence 23, Appl
C 31	44.4	1.4	9347	4	US-10-204-708-36	Sequence 36, Appl
C 32	44.4	1.4	16442	3	US-08-781-891-208	Sequence 208, App
C 33	44.4	1.4	16442	4	US-09-618-166-208	Sequence 208, App
C 34	44.2	1.3	1493	1	US-08-340-820-24	Sequence 24, Appl
C 35	44.2	1.3	1493	1	US-08-593-535-24	Sequence 24, Appl
C 36	44	1.3	240	1	US-08-628-417-6	Sequence 6, Appl
C 37	44	1.3	2394	4	US-09-800-729-33	Sequence 33, Appl
C 38	44	1.3	11050	4	US-10-204-708-85	Sequence 85, Appl
C 39	44	1.3	19233	4	US-10-204-708-46	Sequence 46, Appl
C 40	43.6	1.3	359	4	US-09-621-976-16008	Sequence 16008, A
C 41	43.6	1.3	2915	4	US-09-336-115C-5	Sequence 5, Appl
C 42	43.6	1.3	5501	4	US-10-204-708-38	Sequence 38, Appl
C 43	43.6	1.3	11015	4	US-10-204-708-55	Sequence 55, Appl
C 44	43.4	1.3	140	1	US-08-628-417-5	Sequence 5, Appl
C 45	43.4	1.3	359	4	US-09-621-976-16019	Sequence 16019, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
; Sequence 14, Application US/08232463
; Patent No. 5670367
GENERAL INFORMATION:
; APPLICANT: DORNER, F.
; APPLICANT: SCHRIFFLINGER, F.
; APPLICANT: FALKNER, F. G.
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: Foley & Lardner
; STREET: 1800 Diagonal Road, Suite 500
; CITY: Alexandria
; STATE: VA
; COUNTRY: USA
; ZIP: 22313-0299
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/232,463
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/07/935,313
; FILING DATE:
; APPLICATION NUMBER: EP 91 114 300.6
; FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
; NAME: BENT, Stephen A.
; REGISTRATION NUMBER: 29,768
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-9300
; TELEFAX: (703) 683-4109
; TELEX: 899149
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 7218 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; CLONE: PTZ9PT-F15
; US-08-232-463-14
Query Match 2.6%; Score 85.6; DB 1; Length 7218;

Best Local Similarity 6.4%; Pred. No. 1.6e-12;
Matches 25; Conservative 232; Mismatches 131; Indels 0; Gaps 0;

174 CTTCCTCATTTGCTCTTCTGTCAGTTATTTCTTCTCCGAGCTCGACTCACTA 233
1094 YY 1153
234 CTCTCATCTCCGCGCTTAACTTAAGTCTGCTGCTTACTCTGAATTTCTGC 293
1154 YY 1213
294 CTAGAGCCCTGATGCTCAACCGCATCTGCTGCTGCTGCTGCTGCTGCTG 353
1214 YY 1273
354 TGGAAATATGCTTATGCTTCTGATTTGGAAGTTTGTGCTGCTGCTGCTT 413
1274 YY 1333
414 CCTCATTTTATAGTTCTTGAAGTATGCTGCTGCTGCTGCTGCTGCTGCT 473
1334 YY 1393
474 TTGCTTCTTACCGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 533
1394 YY 1453
534 TTGCTTCTTATGCTTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 561
1454 CTTAACCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1481

RESULT 2

US-09-313-294A-121

Sequence 121, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 121
LENGTH: 272
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc. feature
OTHER INFORMATION: Incyte ID No. 6476212 700548569H1
NAME/KEY: unsure
LOCATION: 14, 32, 61, 127, 162, 192
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-121

Query Match 2.3%; Score 74.4; DB 4; Length 272;

Best Local Similarity 58.7%; Pred. No. 2.4e-10; Indels 3; Gaps 1;

Matches 142; Conservative 0; Mismatches 97;

1777 TATGAGCTTATGAGAGCCATTCCTATGCTCAACAGGCGCATGCTGAGTGA 1836
2 TATGAGCAAGTAAAGCACTATGCTATGCTTCTCTGCGACGCTGATGAGCGTN 61
1837 CTGATGTTTGAAGCACTGCTGCTATTTGAGCGGACGCTCCACCGGAGTTA 1896
62 TTAATATTGAAGCTGAGCTGAGGCTATCATGAGAGCTGATGATTAACCTTT 121
1897 GCTGAGTGGGTTGATGATGATTTGCTGG---GTCAGAAAGCGAGTATGTTTCTGGA 1953
122 GTTANCAAGGTACAGACAGAAATTCATGAGCACTACGANGTTCATTTGCGCTGGT 181

1954 GGTGTTGCCAACTGATGAGCTTCTCTGCAACAGACAGATCTGACATATTCATCA 2013
182 GGGAAAGGACNCTATATGTTTCTTACCAACAAAGAGATATGAGCACTTTAAAG 241
2014 CA 2015
242 CA 243

RESULT 3

US-09-313-294A-463

Sequence 463, Application US/09313294A
Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laligudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
CURRENT FILING DATE: 1999-05-14
NUMBER OF SEQ ID NOS: 7600
SOFTWARE: PERL Program
SEQ ID NO 463
LENGTH: 274
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc. feature
OTHER INFORMATION: Incyte ID No. 6476212 700549278H1
LOCATION: 15, 53
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-463

Query Match 2.2%; Score 72.2; DB 4; Length 274;

Best Local Similarity 60.6%; Pred. No. 9.6e-10; Indels 4; Gaps 2;

Matches 152; Conservative 0; Mismatches 95;

1777 TATGAGCTTATGAGAGCCATTCCTATGCTCAACAGGCGCATGCTGAGTGA 1835
2 TATGAGCAAGTAAAGCACTATGCTATGCTTCTCTGCGACGCTGATGAGCGT 61
1836 TCTGATGTTTGAAGCACTGCTGCTATTTGAGGCGGACGCTCCACCGGAGTT 1895
62 GTTAATATTGAAGCTGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 121
1896 AGCTGATGAGGCTTATGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCT 1952
122 TGTAAATCAAGGTACAGACAGAAATTCATGCGACCTAGCGAAGTTCATTTGCTG 181
1953 AGGTGTTGCGCAACTGATGCTTCTTCTGCAACGAAAGCAAGTCTGCAATTCATCA 2012
182 TGGAAAGGCACTATATGTTTCTTACCAACAAAGAGATATGAGGCACTTTAAACA 241
2013 ACACTCTCAG 2023
242 GCATTCGACG 252

RESULT 4

US-08-232-463-14/c

Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: Foley & Lardner

```

/ STREET: 1800 Diagonal Road, Suite 500
/ CITY: Alexandria
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22313-0299
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/232,463
/ FILING DATE:
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/07/935,313
/ FILING DATE:
/ APPLICATION NUMBER: EP 91 114 300.6
/ FILING DATE: 26-AUG-1991
/ ATTORNEY/AGENT INFORMATION:
/ NAME: BENT, Stephen A.
/ REGISTRATION NUMBER: 29,768
/ REFERENCE/DOCKET NUMBER: 30472/114 IMM
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: (703)836-9300
/ TELEFAX: (703)683-4109
/ TELEX: 899149
/ INFORMATION FOR SEQ ID NO: 14:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 7218 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ IMMEDIATE SOURCE:
/ CLONE: pTZ9pt-Fls
/ US-08-232-463-14

Query Match
Best Local Similarity 9.0%; Score 68.6; DB 1; Length 7218;
Matches 41; Conservative 230; Mismatches 184; Indels 0; Gaps 0;

QY 2731 TTTCGAGAGCTGAGAGAGTGTCAAGCTTCATGATTTTCAGAGAAAGATGAGAGAG 2790
DB 1442 TTTCGACACACACACACACACACACACACACACACACACACACACACACACACAC 1383
QY 2791 TTTCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2850
DB 1382 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1323
QY 2851 AAGAGCATCAGAGAGAGATTTGATCTCGAGAGAAAGATTGATGAGCTTGAGAACAG 2910
DB 1322 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1263
QY 2911 CTCATGTACAGCATGCGCTTCAATGAGATGATGAGACAAAGTCTGATACACAG 2970
DB 1262 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1203
QY 2971 ACAAGACTAAGTTCTTTGTTTCTTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTG 3030
DB 1202 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1143
QY 3031 TCCATTAAATATAGAGCAAACTAAGAGATTAATATATATATATATATATATATATAT 3090
DB 1142 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1083
QY 3091 TAGACGATCTAAGAGACATTAAGTTCTTTGATCTAAGACAAAGTTCTTCTAGATTTT 3150
DB 1082 RRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1023
QY 3151 GTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 3185
DB 1022 GTGAGCGTATGCGAAAGCAAGAAATAATGTTATTA 988

```

```

RESULT 5
US-09-331-581-3
/ Sequence 3, Application US/09331581
/ Patent No. 6130070
/ GENERAL INFORMATION:
/ APPLICANT: TOHDA, Hideki
/ APPLICANT: HAMA, Yuko
/ APPLICANT: KIMAGAI, Hitomichi
/ TITLE OF INVENTION: INDUCTION PROMOTER GENE AND SECRETORY SIGNAL GENE
/ TITLE OF INVENTION: USABLE IN SCHIZOSACCHAROMYCES POMBE, EXPRESSION VECTORS
/ TITLE OF INVENTION: HAVING THE SAME, AND USE THEREOF
/ FILE REFERENCE: 0059-1142-0PCT
/ CURRENT APPLICATION NUMBER: US/09/331,581
/ EARLIER FILING DATE: 1999-06-30
/ EARLIER APPLICATION NUMBER: PCT/JP98/04929
/ EARLIER FILING DATE: 1998-10-30
/ EARLIER APPLICATION NUMBER: JP 9-314608
/ EARLIER FILING DATE: 1997-10-31
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 3
/ LENGTH: 7286
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:DNA
US-09-331-581-3

```

```

Query Match
Best Local Similarity 1.6%; Score 52.6; DB 3; Length 7286;
Matches 97; Conservative 0; Mismatches 74; Indels 0; Gaps 0;

```

```

QY 9 AACAAATTTAAGCAAGTATGTTCTGATCAATTAATTAAGTGAACATTAAGTTA 68
DB 4559 AACAAATCAATGAGAAAGATGTATGTACGTGAAATTAATGACCTCAAGACAAA 4618
QY 69 AGCGAAAAAGAAAAAGTACAAAAATGAAAAATCAAACTGAAATGAAATTT 128
DB 4619 AAAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAGAAAAAG 4678
QY 129 GAGTCCAGATCGAAAAAGAGCGCGTTTGAAGCTTAATTAAGCTTCT 179
DB 4679 TACCTTCTAGGCGGAGAAAGACAGCGGATCAGACATGATTAAGTATACAT 4729

```

```

RESULT 6
US-09-331-581-14
/ Sequence 14, Application US/09331581
/ Patent No. 6130070
/ GENERAL INFORMATION:
/ APPLICANT: TOHDA, Hideki
/ APPLICANT: HAMA, Yuko
/ APPLICANT: KIMAGAI, Hitomichi
/ TITLE OF INVENTION: INDUCTION PROMOTER GENE AND SECRETORY SIGNAL GENE
/ TITLE OF INVENTION: USABLE IN SCHIZOSACCHAROMYCES POMBE, EXPRESSION VECTORS
/ TITLE OF INVENTION: HAVING THE SAME, AND USE THEREOF
/ FILE REFERENCE: 0059-1142-0PCT
/ CURRENT APPLICATION NUMBER: US/09/331,581
/ EARLIER FILING DATE: 1999-06-30
/ EARLIER APPLICATION NUMBER: PCT/JP98/04929
/ EARLIER FILING DATE: 1998-10-30
/ EARLIER APPLICATION NUMBER: JP 9-314608
/ EARLIER FILING DATE: 1997-10-31
/ NUMBER OF SEQ ID NOS: 24
/ SOFTWARE: Patentin Ver. 2.1
/ SEQ ID NO 14
/ LENGTH: 7938
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:DNA
US-09-331-581-14

```


Db 15736 TTAATTTAATAATTTTTCTTTTTTTTGTGTTTATGACATATAATTTTTTTTTTT 15677

Qy 617 TTAATCTCAGCTCATGTCGGAATTTTGAGATTTGGTAGACCTGGGTTCTT 675

Db 15676 AATGTTTTTTTTCTCTTGTTGTTTATTTTTTAATAACATTTTTTTTTTATAT 15618

RESULT 12

```

US-10-204-708-69
/ Sequence 69, Application US/10204708
/ Patent No. 6677731
/ GENERAL INFORMATION:
/ APPLICANT: OLEK, Alexander
/ APPLICANT: PIEPENROCK, Christian
/ APPLICANT: BERLIN, Kurt
/ TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
/ TITLE OF INVENTION: by Assessing DNA Methylation
/ FILE REFERENCE: 5013.1012
/ CURRENT APPLICATION NUMBER: US/10/204,708
/ CURRENT FILING DATE: 2003-05-06
/ PRIOR APPLICATION NUMBER: PCT/EP01/03971
/ PRIOR FILING DATE: 2001-04-06
/ PRIOR APPLICATION NUMBER: DE 10019058.8
/ PRIOR FILING DATE: 2000-04-06
/ PRIOR APPLICATION NUMBER: DE 10019173.8
/ PRIOR FILING DATE: 2000-04-07
/ PRIOR APPLICATION NUMBER: DE 10032529.7
/ PRIOR FILING DATE: 2000-06-30
/ PRIOR APPLICATION NUMBER: DE 10043826.1
/ PRIOR FILING DATE: 2000-09-01
/ NUMBER OF SEQ ID NOS: 98
/ SEQ ID NO 69
/ LENGTH: 6040
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-69

```

Query Match	1.5%	Score 49.2	DB 4	length 6040
Best Local Similarity	49.5%	Pred. No. 0.0094		
Matches 186	Conservative 0	Mismatches 183	Indels 7	Gaps 2

QY	ATTTCGAGGTTTTTGTCGACGAGTTACTTTTCCCTAATTTTATAGTCTTAGGTA	438
Db	ATTGTTAGTGGTTTTCGTTTTTATTTATGTTAATTAAGTATTATTAATAAATA	3397
QY	ACGATACCTGGCTCTTACTGTTTTGTTCATTTTGTGAGCTTACCGCTTAACTGGTG	498
Db	ATTAAATTATTAAGTTTATTTTATTTTATAGTATCTGTGTTTTATGCAATTAAGTAGT	345

QY	499	A-----TCGAGTATTGACTGAGAAAAATCCCTCGTTTTCGTTTGGTTTGGTATATATAA	554
Db	3458	AAGATTTCGGGATTTGGATTATATAAAGATTGAGTATCGTTTATATATATATATAGG	3512
QY	555	TCGATTGATACCTTTTGTCCTTGAGTTGGTTTTCAGCCTAAGCGTTGTGGCT	614
Db	3518	TGTCATTGAGAGTAAGTAGTTTTTTTATGGTTTATTTTTTTT---TATTGGAATATATT	3574

OY	615	TGTTAAACTTCACGTCATGTCGATTTTGAGATTTTGGTAGCACTGCGGTTCTT	674
Db	3575	ATTAGAACTTGTTGTGAATTTGATTTTATGTTTAAATGATTTTGGAAAAAG	3634
OY	675	TGTCGGCATATAGTTGTAATAATGATGCTCTGCGGCATCTCAGGAAAAAGAC	734
Db	3635	TTTGCTATATAGAGAGCTGATATATATATAAGATTTTGTATTTTGGAAAAATAGAT	3699

RESULT 13

US-08-781-891-209/c
Sequence 209, Application US/08781891
; Patent No. 6090620
; GENERAL INFORMATION:
Inventor: David W. ...

	Query Match	1.5%	Score 48.6;	DB 3;	Length 51259;
	Best Local Similarity	52.2%	Pred. No. 0.048;		
	Matches 108;	Conservative	Mismatches 99;	Indels 0;	Gaps 0
Qy	2764	GAGTTCAAGGAAAGAGATGAGAGAGTTTGTGAGAGAGAGAGATCTGTAAGAAT	282		
Db	232	GAGAGAGAGAGAGAGAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG	173		

[illegible]

RESULT 14

US-09-618-166-209/c
; Sequence 209, Application US/09618166
; Patent No. 6583112
; GENERAL INFORMATION:
; APPLICANT: Fu, Ying-Hui
; Yu, Chang-En
; Oshima, Junko
; Mulligan, John T.
;

schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSEE: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052,419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 209:
US-09-618-166-209
Query Match 1.5%; Score 48.6; DB 4; Length 51259;
Best Local Similarity 52.2%; Pred. No. 0.048;
Matches 108; Conservative 0; Mismatches 99; Indels 0; Gaps 0;
QY 2764 GAGTTTCAGAGAAAGATGATGAGATTGTGGAAGAGAGAGATGCTGATPAAAGAT 2823
DB 2332 GAGGAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 173
QY 2824 CAGAGAGAGAGATGAGAGCATGAGAGAGAGGAGCATGAGAGATTTTGTCTGGAG 2883
DB 172 GAGCAG 113
QY 2884 AAGAAATTTGATGAGGCTTTGGAACAGCTCATGTATACAGCATGAGCTTCAATGAAGAT 2943
DB 112 AAGGAGGAGGAG 53
QY 2944 GATTGAGACAAAAGCTGTGTAACAAG 2970
DB 52 GAGGAG 26
RESULT 15
US-09-014-969-14/c
Sequence 14, Application US/09014969
Patent No. 5965397
GENERAL INFORMATION:
APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallie, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Merberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32

CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2447 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-014-969-14
Query Match 1.5%; Score 47.8; DB 2; Length 2447;
Best Local Similarity 47.2%; Pred. No. 0.013;
Matches 142; Conservative 1; Mismatches 158; Indels 0; Gaps 0;
QY 390 TTTTGCTATGAGTACTTTTCCCTATATTTATAGTCTTGTAGTAACGATACCTGC 449
DB 2447 TTTTGTCTATGAGTACTTTTCCCTATATTTATAGTCTTGTAGTAACGATACCTGC 2388
QY 450 GTCTTACGTTTGTGTCATTTGTGTGCTTACACCGTTTGTAGTGTGAGATATT 509
DB 2387 TTTTGTCTATGAGTACTTTTCCCTATATTTATAGTCTTGTAGTAACGATACCTGC 2328
QY 510 TGACGTGAAATCCCTGCTTTTGTGTTGTTCAATATAATCGATTGATCTACC 569
DB 2327 TTTTGTCTATGAGTACTTTTCCCTATATTTATAGTCTTGTAGTAACGATACCTGC 2268
QY 570 TTTTGCTTTGATGTTTGTGTTTGTAGGCTATGCGTTGTGCTGTATTAATCTCAG 629
DB 2267 TTTTGTCTTTGATGTTTGTGTTTGTAGGCTATGCGTTGTGCTGTATTAATCTCAG 2208
QY 630 TTGATGTGTGATTTTGAATTTTGTGTAGTACTGTGGGTTTCTTTGGTGCTATAGTT 689
DB 2207 ATCAATATACAGATCCAGATATGTGAACATATATATATATATATATATATATAT 2148
QY 690 G 690
DB 2147 G 2147

Search completed: April 8, 2004, 16:29:38
Job time : 170.206 secs

This Page Blank (uspto)


```

QY 1269 CAAAAGAGCGATGATCAAGAAAGAGATAGTGTCAAAAAGTTCTTTGGCAGCTTG 1328
Db 2033 CAAAAGAACGAGAGACTGTAAAGAAAGCAATGTGTTAAAGATTTCTTTGAAGACTTG 1974
QY 1329 GATAGCTTGTGATGAGAGATTAATTAACACAGAGAGCGATGCTTGTCCAGCTGT 1388
Db 1973 GATGAGCTTATCTATCGAAAGATCAATGAACCGAAAGGCGATGGCACTGTACAGCTTGT 1914
QY 1389 CAGAACGACCTGTGTCATGATTTGTATA---ACTGACCTCTTACTAGCTTCATCCG 1445
Db 1913 CGAGGTGTCTCCGTCTCTATTTGTTGTACAGAGAGCTGACCTCTGTATTAATCTATCC 1854
QY 1446 AGGACAAAAGAGAGCTAGGCGAGTTAAGCTCCATAGAGATTTGCTGAAGTTTAAAGAA 1505
Db 1853 AAAACAAAGGGGTCAAAAAGGGTGAAGATCCATAGGAGATTTGCTGAGCTTTTGAAGAG 1794
QY 1506 GATCTACAGATGAGAGCGGCTGTCTGATTCCTTGTGTGATTTATGAGCAATGAGAG 1565
Db 1793 GAATCGCGCAGAAAGGGCTCTGCAATTAATCCAGCTGGGGAAGTATTTGTAGTGAAGAA 1734
QY 1566 GATTGGGTGAGATGAAAGATTAATGAAATTTGTCTGGCTCCATGGTCAATCATG 1625
Db 1733 GATTAAAGATGAGAGAAAGATCATGAATAGTTGGCTCCGATGGTTGTCATTTCA 1674
QY 1626 AATCTAGACTGTATGAGAGATPAAGATGAGTGAATTTCTGTCTTTTACTTCTT 1685
Db 1673 AATCAAGGCTTGAACGATGAAATGATA----- 1643
QY 1686 TAATTTTCTCTTCATTTCTACTGATCTTGAATGTTACATGTAGTGGTCCGCAATGG 1745
Db 1642 -----AGTGGCTAGGTATGGG 1627
QY 1746 CAACCAAGAGCTCTGGAATCTTCAACAAGTATGAGGCTTTAGACAGCCCAATTTCTTA 1805
Db 1626 TAACCAAGAGCTCTTCAATTAATTTTATGATGATGATGCTGTGAAGCTGACAGCGCTTA 1567
QY 1806 TGGTCACAGGAGCATCTGTGGATAGTGTCTGATGTTTGAAGCAGTGCCTGCTA 1865
Db 1566 TGGCCCCAGGCTATCCGGGGTGTAGGCTTTGATTTGAACATCAAGCTATAGGTTA 1507
QY 1866 TTTGAGGCGGAACCGCTCCACCGGAGTTAGCTGAGATGGGGTTAGATAGATTTGCTG 1925
Db 1506 TCTTGGGCTGAGCGCTTACAGAACACTTTGCAAGAACAGAACTGACGAGATGCTG 1447
QY 1926 GGGTCAGAGCGGATGATTTT-----CTGAGGTGTTGCCAATGTTATGCTTCT 1979
Db 1446 GTTTATGTCACACCGTATGATTTTCTCCGTGGCGGCGAGCTCTATGATACAT 1387
QY 1980 TGCACGAGAGAGATGTGAGATTAATCAATCAACAATCTCAAGGT 2025
Db 1386 GGCATATAAAGAGACCTGACTTTTCAACAGACATTTCCAAAGT 1341

```

RESULT 2
US-10-424-599-91177

Sequence 91177, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 91177
TYPE: DNA
ORGANISM: Glycine max

```

FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_53343C.1
US-10-424-599-91177
Query Match
Best Local Similarity 7.7%; Score 253; DB 12; Length 1377;
Matches 428; Conservative 64.8%; Pred. No. 1.8e-53;
Indels 13; Gaps 3;
Mismatches 220;
QY 1047 TATGATTAACAATTTGTGTGCAACCCCACTGTATCTCCCTCTTTGAAAGAGATGG 1106
Db 29 TATGAAGAAGACTATGTGATCTCAATCTGTATTTGGGCTCCATCGACATGCGCTGG 88
QY 1107 AATTGGCAGGCAAGAGAGATTTCTGTCAGACACAGCTGTGACAGATTTCTGAGCTG 1166
Db 89 AACTGGCAATCCAGACTGTGGCAATTCAGTCCAAATGTAAGGATGAATTTACCAAGAG 148
QY 1167 G-----AGATGATGTGATTAATGCTTCTGAGGAAGAGATTTCCAGTCTTTG 1217
Db 149 GACCTCAAAAAGAAATTAATGTTGTTGATGATGATGGGAGAGAGAAATCTGATATTTG 208
QY 1218 GATGATTCGATGACGACCTTGCAAGTATGATTAATGATCTGAGATGTGATCAAAAAGAC 1277
Db 209 GAAGATCTGATGATGATGATCTGATGATGATGATGATGATGATGATGATGATGATGAT 268
QY 1278 CATGATCAAGAAAGCAATTAATGTTGTTCAAAAAGTTCTTTGGCAGCTTGGATAGCTTG 1337
Db 269 CATGAGACAGTAAAGAAAGCAATGATTAAGATTTCTTGAATTTTGAATGGCTTG 328
QY 1338 TCGATCAAGCAGTAATTAATGAACCAAGAGCAGTGGCATTTGTCAGCTTGTCAAAACGA 1397
Db 329 ACTGTTGAACAGATTAACGAACCAAGAAAGAGATGGCATTTGTCAAGCTTCCAAAGTGGT 388
QY 1398 CCGTGTGCGCATGATGTTGTATTAAC---CTGACCTCTTACTACTCATGAGGAGCAAAA 1454
Db 389 CCGAGAGTATGATTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 448
QY 1455 GAGCTAGCGAGTTAAGCTCCATAGAGATTTGCTGAAGTTTGAAGAGATCTACAG 1514
Db 449 GATCAAAAAGGGGTAAATTAATCAAGGAGCTTGTCTATCTTTTGAATGAGAAATTAAGC 508
QY 1515 ATGAGAGGCGCATCTGTCTCTTGTGTTGATGATTTATGAGCAGTGGAGGCTTGGGT 1574
Db 509 AAAAGGGTACTTCAATTAATTCAGCTGGGGAAGATTTGAAAGTGAAGGTTGAAA 568
QY 1575 GAGATGAAGAAGATTAATTTGCTGGCTCCCAATGTCATCATCATGATTAATCTAGA 1634
Db 569 GAAGAG-AAAAGACCATGATGATTTGGCTCCAAATGTTGTCATCAAGATTAAGAA 627
QY 1635 CTGATTAAGACGATTAAGATTAAGTGAATTTCTGCTTTTACTTTTAAATTTTC 1694
Db 628 CTGAAACAGATGAATAATGATTAAGCGATTAATCATTAATTAATTTTACTTATTAAT 687
QY 1695 T 1695
Db 688 T 688

```

RESULT 3

US-10-424-599-91176
Sequence 91176, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 91176

LENGTH: 1589
 TYPE: DNA
 ORGANISM: Glycine max
 FEATURE:
 OTHER INFORMATION: Clone ID: PAT_MRT3847_53342C.1
 US-10-424-599-91176

Query Match 6.6%; Score 215.6; DB 12; Length 1589;
 Best Local Similarity 64.6%; Pred. No. 6.9e-44; Indels 12; Gaps 2;
 Matches 358; Conservative 0; Mismatches 184

1047 TATGATTAACACTTTGTGGACCCCACTGTATCTCGCCCTCTTGGAAAGAGATGG 1106
 29 TATGAAGACACTATGTGACTCCAAATCTGTATTCGGCTCCACGTGAGATGGCTGG 88
 1107 AATTGGCAGCAAGAGAGAGTTCTGCTCAGACACAGCTGTGAGAGATTCTTGACGTG 1166
 89 AACTGGCAATCCAGACCTGTGTGCAATTCAGTCCAAATGTAAAGGATGAAATTTTCCAGAG 148
 1167 G-----AGGATGATGTGATTAATGCTTCTGAGAAAGAAATGATTCGATGCTTG 1217
 149 GACCTTCAAAAGATTAATGTTGATGATGATGAGGAGAGAGAGATCTGATGATTTG 208
 1218 GATGATTTGATGACGACCTTGGCAAGTATGATTAATGATCTCGATGTGATGATGATGAGC 1277
 209 GAAGATAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
 1278 CATGATCAGCAAGAAAGCAATTAAGTGTCTCAAAAGTTCTTTGGCAGCTTGTGATGCTTG 1337
 269 CATGACACAGTAAAGAAAGCAATTAAGTGTCTCAAAAGTTCTTTGGCAGCTTGTGATGCTTG 328
 1338 TCGATGACACAGTAAATGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1397
 329 ACTGTTGAACAGTAAAGCAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 388
 1398 CCTGTGCACTGATTTGATTA---ACCTGCACTCTCTACTAGCTGATGATGATGATGATGATG 1454
 389 CCGGAGCTATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
 1455 GGAAGCTAGGAGATTAAGCTCCATTAAGAAATTTGCTGAAAGTTTGAAGAAAGATCTACAG 1514
 449 GGATCAAAAGAGGCTGAAGATTCACAGGAGCTTGTATCTTTGATGATGAGAAATTAACGC 508
 1515 ATGAGAGGGGCACTGCTGATTCCTTGTGATGATTAATGAGGAGTGAAGGCTTTGGGT 1574
 509 AGAAGGGGCTCTGCAATTAATCCACTGCGAAGTATTGTAAGTGAAGCGTTTAAAG 568
 1575 GAGATGAAAGAGA 1588
 569 GTGCAGCAGAAAGA 582

RESULT 4
 US-10-425-114-18351
 ; Sequence 18351, Application US/10425114
 ; Publication No. US20040034888A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Liu, Jindong
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Kovalic, David K.
 ; APPLICANT: Screen, Steven E.
 ; APPLICANT: Tabaka, Jack E
 ; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCES: 38-21(5313) B
 ; CURRENT APPLICATION NUMBER: US/10/425,114
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 73128
 ; SEQ ID NO 18351
 ; LENGTH: 2205
 ; TYPE: DNA
 ; ORGANISM: Zea mays

FEATURE:
 OTHER INFORMATION: Clone ID: LIB3079-030-A8_FLI
 US-10-425-114-18351

Query Match 6.0%; Score 198; DB 12; Length 2205;
 Best Local Similarity 55.7%; Pred. No. 2.6e-39;
 Matches 518; Conservative 0; Mismatches 330; Indels 82; Gaps 4;

1157 TCTGACGCGAGAGATGATGATTAATGCTTCTGAGAGAGAGAAATTCGATGCTTT 1216
 452 TTCTGTCCCTGAGGTGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 511
 1217 GATGATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1276
 512 TGAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 571
 1277 CCATGATCAGCAAGAAAGCAATTAAGTGTTCAAAAGTTCTTTGGCAGCTTGTGATGCTT 1336
 572 TTTTGAGACTCGAAAAAGCAAGAGGTTCAAGAGTTCTTTGAAGTATTGATTCATT 631
 1337 GTGATGACAGATTAATGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1396
 632 GAGTTTGAACAAATTAATGAACAACTGGCAATGATGATGATGATGATGATGATGATGATG 691
 1397 ACTGTGCTCATGATTTGATTA---CCTGCACTCTACTAGCTGATGATGATGATGATGATG 1453
 692 ACTGTGCTCATGATTTGATTA---CCTGCACTCTACTAGCTGATGATGATGATGATGATG 751
 1454 AGAGCTAGAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1513
 752 GGGTTCTCAAGAGGTTAACTTCAAGAGATTTGGTGTGATTTGATGATGATGATGATGATG 811
 1514 GATGAGAGCGCATCTGTATCTTTGTGATGATTAATGAGAGAGAGAGAGAGAGAGAGAGAG 1573
 812 TCGAGGGGCACTTCAAGTTTACAGCGGGTAAACAAATTTGGAAATGAGAAAGAGCTG-- 869
 1574 TGAGATGAAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1633
 870 -CAAGAAAGCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 928
 1634 ACTGATTAAGAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAGATTAAG 1693
 929 CTGGAAGAAAGATGAAGATGATA----- 951
 1694 CTCTTGATTTCAATGATCTTAAGATTTCAATGATGATGATGATGATGATGATGATGATG 1753
 952 -----AGTGAAGGGCATGGAAACCAAG 975
 1754 AGCTGTGATTAATCTTCAAGATTAAGAGCTCTTAAGACAGAGCTCTTAATGATGATGATG 1813
 976 AGCTCTTATTAATTTTGAAGATTAAGAGCAAGTAAAGCAAGTAAAGCAAGTAAAGCAAGT 1035
 1814 AGGGCCATGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1873
 1036 CTGGGACAGCTGTATGATGATGATGATTAATTTGAAAGCTCAGCTGAGGCTATGATGATG 1095
 1874 CCGAAGCTTCAACCGGAGATTAAGCTGATGATGATGATGATGATGATGATGATGATGATG 1932
 1096 CTGAAGCTGTGATTAACCTTTGTTAAACAAAGTAAACAAAGTAAACAAAGTAAACAAAGT 1155
 1933 --AAGGAGATTAATTTTGTGAGAGTTGCGCACTGATGATGATGATGATGATGATGATG 1990
 1156 GCAAGTTGATTTTGTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1215
 1991 AAGATCTGACATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2050
 1216 AGGATATGAGGACATTAACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 1275
 2051 ATGCTTTTATTTTGTGATGATGATTA 2080
 1276 GGTCTTACATGATGATGATGATGATTA 1305

RESULT 5
US-10-424-599-73594/c
; Sequence 73594, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 73594
; LENGTH: 2890
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37470C.1
US-10-424-599-73594

Query Match 6.0%; Score 197.8; DB 12; Length 2890;
Best Local Similarity 69.4%; Pred. No. 3.6e-39;
Matches 284; Conservative 0; Mismatches 122; Indels 3; Gaps 1;
QY 1243 GTGATGATTATGACTCGGATGTGTCATAAAGCCATGATCCGAAAGCAATTAAT 1302
DB 2890 GTGATGATTATGACTCGGATGTGTCATAAAGCCATGATCCGAAAGCAATTAAT 1302
QY 1303 GGTTCAAAAAGTTCTTTGGCAGCTTGATAGCTTGATCGATCGACAGATAATGAACAC 1362
DB 2830 GGTTCAAAAAGTTCTTTGGCAGCTTGATAGCTTGATCGATCGACAGATAATGAACAC 1362
QY 1363 AGAGCAGTGGCAGTGGCAGCTTGATAGCTTGATCGATCGACAGATAATGAACAC 1419
DB 2770 AAAGCAGTGGCAGTGGCAGCTTGATAGCTTGATCGATCGACAGATAATGAACAC 1419
QY 1420 ACCTGACCCCTCTACTAGCTCATGCGAGACAAAGAGGATAGCGAGTTAACTCCATA 1479
DB 2710 GACTGACCCCTCTACTAGCTCATGCGAGACAAAGAGGATAGCGAGTTAACTCCATA 1479
QY 1480 GAGAAATGGCTGAAGTTTGAAGAGATCTACAGATGAGAGGCGCATCTGTCTTCTT 1539
DB 2650 GGAAGTTTGGCTGAGCTTTTGAAGAGATCTACAGATGAGAGGCGCATCTGTCTTCTT 1539
QY 1540 GTGATGATTATGAGGAGTGGAGGAGGTTGGTGAAGATGAAGATTAATGAATG 1599
DB 2590 CTGGGAGAGTATTTGTAAGTGAAGGTTTAAAGATGAGAGATCATGAATAG 2531
QY 1600 TCTGGCTCCATGCTCATCATCATGAATTAAGCTGTGTAAGACGA 1648
DB 2530 TTGGCTTCTGATGTTGATCATCAATACAGAGCTTGAACAGCCTGA 2482

RESULT 6
US-10-424-599-73591/c
; Sequence 73591, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 73591
; LENGTH: 1375

TYPE: DNA
ORGANISM: Glycine max
FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1375)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: PAT_MRT3847_37468C.1
US-10-424-599-73591

Query Match 5.2%; Score 169.8; DB 12; Length 1375;
Best Local Similarity 69.0%; Pred. No. 3e-32;
Matches 247; Conservative 0; Mismatches 108; Indels 3; Gaps 1;
QY 1243 GTGATGATTATGACTCGGATGTGTCATAAAGCCATGATCCGAAAGCAATTAAT 1302
DB 1375 GTGATGATTATGACTCGGATGTGTCATAAAGCCATGATCCGAAAGCAATTAAT 1316
QY 1303 GGTTCAAAAAGTTCTTTGGCAGCTTGATAGCTTGATCGATCGACAGATAATGAACAC 1362
DB 1315 GGTTCAAAAAGTTCTTTGGCAGCTTGATAGCTTGATCGATCGACAGATAATGAACAC 1362
QY 1363 AGAGCAGTGGCAGTGGCAGCTTGATAGCTTGATCGATCGACAGATAATGAACAC 1419
DB 1255 AAAGCAGTGGCAGTGGCAGCTTGATAGCTTGATCGATCGACAGATAATGAACAC 1419
QY 1420 ACCTGACCCCTCTACTAGCTCATGCGAGACAAAGAGGATAGCGAGTTAACTCCATA 1479
DB 1195 GACTGACCCCTCTACTAGCTCATGCGAGACAAAGAGGATAGCGAGTTAACTCCATA 1479
QY 1480 GAGAAATGGCTGAAGTTTGAAGAGATCTACAGATGAGAGGCGCATCTGTCTTCTT 1539
DB 1135 GGAAGTTTGGCTGAGCTTTTGAAGAGATCTACAGATGAGAGGCGCATCTGTCTTCTT 1539
QY 1540 GTGATGATTATGAGGAGTGGAGGAGGTTGGTGAAGATGAAGATTAATGAAT 1597
DB 1075 CTGGGAGAGTATTTGTAAGTGAAGGTTTAAAGATGAGAGGATCATGAAT 1018

RESULT 7
US-10-424-599-36215/c
; Sequence 36215, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 36215
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
NAME/KEY: unsure
LOCATION: (1)..(1103)
OTHER INFORMATION: unsure at all n locations
OTHER INFORMATION: Clone ID: PAT_MRT3847_132704C.1
US-10-424-599-36215

Query Match 3.2%; Score 103.8; DB 12; Length 1103;
Best Local Similarity 67.3%; Pred. No. 1.7e-15;
Matches 175; Conservative 0; Mismatches 83; Indels 2; Gaps 2;
QY 2126 ATTCCATGAGCAAAACAGGCTGAATTTGAGTTGAATATCAACAGAG-ATGTTGTA 2184
DB 359 ATTCCATGAGTAATCTTAAGCTCAATATGACATATGATCAGAGAAATGTTGTA 300

QY 2185 AAGAGCTGAGGAGATCTCTGAGCAATCAGCAGTGAATCTTAAAGCAAGCTC 2244
DB 2299 CACCAATTCGGGAATGATGAGACCAACCAAGCTGCTCTATTTTAAAGCAAGCT 240
QY 2245 TCAAAACAGAACAGACGCCAAGGTCCTTGAAGATCTCTGAAATTTATGACGAGAG 2304
DB 2339 GTCAAAAGTGCAAAACACAAAAGCTCTGAAAGATCTATTTAGTATTAATGCTGAGAG 180
QY 2305 CTGCGTGAAGCTGACGAGATATCGATGCTGAGACA-GAGAACTAAGTGAAGCATGA 2363
DB 179 CTGCGAAGCAATGAGGAAATTCGATGTGAGGAGAGAACTTAAATGCAACATGA 120
QY 2364 ACAGAACAGGAGAGATAT 2383
DB 119 AGAGACCAAAAAGNAGATGT 100

RESULT 8
US-10-312-841-2
; Sequence 2, Application US/10312841
; Publication No. US20030186277A1
; GENERAL INFORMATION:
; APPLICANT: Epigenomics AG
; TITLE OF INVENTION: Diagnose von bededeutenden genetischen Parametern innerhalb des MHC
; FILE REFERENCE: E01/1208/WO
; CURRENT APPLICATION NUMBER: US/10/312,841
; CURRENT FILING DATE: 2002-12-30
; NUMBER OF SEQ ID NOS: 2
; SEQ ID NO 2
; LENGTH: 3673778
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (379615)
US-10-312-841-2

Query Match 2.5%; Score 81.4; DB 14; Length 3673778;
Best Local Similarity 45.6%; Pred. No. 2.2e-07;
Matches 286; Conservative 0; Mismatches 341; Indels 0; Gaps 0;
QY 80 AAAAAAAGTCAAAAATGAAACAAATCAAACTGAATGAATTTGAGTCCAGAA 139
DB 971449 AAAAAAAGTCAAAAATGAAACAAATCAAACTGAATGAATTTGAGTCCAGAA 139
QY 140 TCGGAAAAAGAGCCGTTTATAGCTTAATAAGTTCCTCATTTGCTCTCTGCTCA 199
DB 971509 AAAGTAATTAAGAAATAGTTTACGGTTTTTTTTTTTTTTTTTTTTTTTTTTT 971508
QY 200 GTTATTTTCTCTCCGAGGCTCTGACCTGCTGCTCCGCGGCTTAAACTT 259
DB 971569 TTTTCT 971568
QY 260 AGCTTCGCTGCTTACTCTGTAAGTTTCTGCTTGAAGCTCGATGCGCTCACCGC 319
DB 971629 TTTTCT 971628
QY 320 ATGCAATCTGCTCGATTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCA 379
DB 971689 TTTTCT 971688
QY 380 TTTCGAAGTTTGTGCTATGGGTTACTTTTCCCTATATTATTAAGTCTTAGGTA 439
DB 971749 TTTTCT 971748
QY 440 CGATACCTGCTGCT 499
DB 971809 TTTTCT 971808
QY 500 TCGGAGTATTTGACTGTGAAAAATCCTGCTTTTGTGTTTGTTCATATATAATCGGA 559

DB 971869 TTTTCT 971868
QY 560 TTGATCTACTTTTGTGCTTGTGATGTTTGTGTTTGTGAGCTTATGCTTGTGCTGTA 619
DB 971929 TTTTCT 971928
QY 620 TAACCTCAGCTTCATGTCGATTTTGAATTTTGTGTAAGTCTGTCGCTTCTTGTG 679
DB 971989 TTTTCT 971988
QY 680 GCTATAGTTTGAATAATGATTTCTTG 706
DB 972049 AGCTTGAGTGTAGTGTATTTAGT 972075

RESULT 9
US-10-424-599-99770/c
; Sequence 99770, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovallig David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223) B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 99770
; LENGTH: 951
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_61108C.1
US-10-424-599-99770

Query Match 2.4%; Score 77; DB 12; Length 951;
Best Local Similarity 63.0%; Pred. No. 1.1e-08;
Matches 119; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 2736 CAGAGCTGAGAGTGTCAAGCTTCATCGAGTTTCAAGAGAAAGATGAGAGATTGT 2795
DB 606 CAGGTGGAATATCTTCTGATGTTTCTGAGCTCAAGCAAGAGATGAGCAATTGCA 547
QY 2796 GGAAGAGAGAGATGCTGATTAAGATCAAGAGAAAGATGAGAGATGAGAGAGAG 2855
DB 546 GCGAGAGAGAGAGAAATTAATAAATTCATGAGAGAGAAAGTTGCTGAGAGAGAA 487
QY 2856 GCATCAGAGAGATTTGATCTGAGAGAAAGATTTGATGAGCTTTGAAACAGTCA 2915
DB 486 GCGAGTGGCAAGAGAGGTTGAACCTGAGAGAGAGATTGAGAAATGAACTGACGACCTCAT 427
QY 2916 GTACAAAGCA 2924
DB 426 GGCACAGTA 418

RESULT 10
US-10-311-455-1670
; Sequence 1670, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determi
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16

RESULT 11
US-10-239-676-52
; Sequence 52, Application US/10239676
; Publication No. US20030082609A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIRENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
; FILE REFERENCE: 5013.1003
; CURRENT APPLICATION NUMBER: US/10/239,676
; CURRENT FILING DATE: 2002-09-24
; PRIORITY APPLICATION NUMBER: PCT/EP01/03968

Oy 251 TTAAACTACGCTCCGCGCTCCTTACCTCCTGAAGATTTCGCGCTTAGACCTCCGATCG 310
 Db 202 TTTTATTTTGTATTTTTCGTTTATTTTGTGATTTTGTGCGTTTTTTTTTTTATTTTTTTTC 261
 Oy 311 CCTCACCGCATGATCTCTGCTCGATTTCTCTTTTCTTCGCTGAAAAATTGCCCTA 370
 Db 262 GTTTTTTTTTTTTGTGCGTTGCTGTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTATTT 321
 Oy 371 TGTTCGATTTGGAAGTTTTTGTGCTATGAGGCTACTTTTTCGCCATAATTTATAGT 430
 Db 322 TTATTTGTTTTTATTTTTTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTT 381
 Oy 431 CTTAGTAACGATACCGCGCTCTACTCTGTGTTTGTCAATTTGTGTGCTTACCGTT 490
 Db 382 TTTTTTTGTTTTTTTTTGTGTTTTTTTTTTTTTTTTTATTTTTTTTTTGTGTTTTTTTT 441
 Oy 491 AGTCGCTAGTCGAGATTGTGACTGCTGAAAAAACCCTGTTTTTGTGTTTGTTCATA 550
 Db 442 ATTTTTTTTTTTTTTTTATTTTTTTTTTTTATTTTTTTTTTGTGTTTTTTTTTTATTT 501
 Oy 551 TAAATCGATGATCTACCTTTTGTGCGTTGATGTTGTTTTTGAACCTATGCGTTT 610
 Db 502 TTTTTTTTTTTTTTGAATTTTTTTTTTTTTTTTTTTTTTGTGTTTTTTTTTGTTTTT 561
 Oy 611 GCGCTGTAACTTCACTCACTGCTCAAGCTGATTTTGAATTTTGTGAGAGTCACTGCGTT 670
 Db 562 TTTTTTTTTTGTATTTTATTTTTTTTTTTTTTTTATTTTTTTTATATTTTATTTTTTTTT 621
 Oy 671 TCTTTGTGCGTATGATGTTTAAATAGCTTCT 704
 Db 622 TTATTTATATATATTTTTTTTTTTGAGTAGGTTT 655

APPLICANT: BRELIN, KYLE
 TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA
 TITLE OF INVENTION: Transcription
 TITLE OF INVENTION: by Means of Assessing the Methylation Status of Genes Associated
 TITLE OF INVENTION: with DNA Transcription
 FILE REFERENCE: 5013.1009
 CURRENT APPLICATION NUMBER: US/10/240,453
 CURRENT FILING DATE: 2002-10-02
 PRIOR APPLICATION NUMBER: PCT/EP01/03973.
 PRIOR FILING DATE: 2001-04-06
 PRIOR APPLICATION NUMBER: DE 10019058.8


```

1 RESULT 13
2 US-10-424-599-57885
3 : Sequence 57885, Application US/10424599
4 : Publication No, US20040031072A1
5 : GENERAL INFORMATION:
6 :
7 : APPLICANT: La Rosa Thomas J
8 : APPLICANT: Kovalic David K
9 : APPLICANT: Zhou Yihua
10 : APPLICANT: Cao Yongwei
11 :
12 : TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
13 : TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
14 :
15 : FILE REFERENCE: 38-21(53223)B
16 :
17 : CURRENT APPLICATION NUMBER: US/10/424,599
18 :
19 : CURRENT FILING DATE: 2003-04-28
20 :
21 : NUMBER OF SEQ ID NOS: 285684
22 :
23 : SEQ ID NO 57885
24 :
25 : LENGTH: 673
26 :
27 : TYPE: DNA
28 :
29 : ORGANISM: Glycine max
30 :
31 : FEATURE:

```

```

RESULT 14
US-10-311-455-1931
; Sequence 1931, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Methylation of Cytosine Residues in DNA
; TITLE OF INVENTION: Cytosine methylation
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCT/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1931
; LENGTH: 14006
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
; FEATURE:
; NAME/KEY: unsure
; LOCATION: 8289..8310..8313
; OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1931
Query Match 2.0%; Score 66; DB 14; Length 14006;

```



```

QY 336 ATTCTCTTTTCTGCTGGAATAATGCCCTAATGTTCTCGATTTCGAAGGTTTTGT 395
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2922 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 2981
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 396 GCTATGGGTACTTTTCCCTATATTTATAGTCTTAGTAACATACCGCTTA 455
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 2982 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3041
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 456 CTGTTTGTGTAATTTGTTGCTTTCACCGTTTATGTCGATGAGATTGACATG 515
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3042 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3101
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 516 TGAATAATCCCTGCTTTTGTGTTGTTTCAATATAATGAGATTGATCTTATG 575
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3102 TGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3161
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 576 GCTTGAATGTTGTTTGTGAGCCATAGCGTGTGCTGTATTAACCTACGTTTCA 635
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3162 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3221
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 636 TGTGATTTTGAATTTTGTAGTACGTCGAGTTCTT 675
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB 3222 TTTTGTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 3261
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

RESULT 3

ABV40163/C
ABV40163 standard; cDNA; 556 BP.

AC ABV40163;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 40154.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KM pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.

PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

PD WPI; 2001-662795/76.

PS Claim 1; Page 8115; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising

CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the

CC specification or its complement. (I) is useful for: (a) assessing whether

CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)

CC determining whether prostate cancer has metastasized in a patient; (h)

CC assessing the aggressiveness or indolence of prostate cancer in a patient

CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

QY Query Match 9.9%; Score 68.6; DB 5; Length 556;

DB Best local Similarity 48.3%; Pred. No. 8.6e-06;

DB Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

QY 265 CTCCGCTGTTACTCTGTAAGTTTCTGCTTAGAGCTTCGATGCGCTACCGCATGCA 324

DB 486 CCCCCCTCCCTTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 427

QY 325 TTCTGCTGCATTTCTCTTTTCTGCTGGAATAATGCCCTAATGTTCTGATTTG 384

DB 426 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 367

QY 385 AAGGTTTTGTGCTAAGGTTACTTTTCCCTAATTTTATGTTCTTAGGTAACGATA 444

DB 366 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 307

QY 445 CCTGCGCTTACTGTTTGTTCATTTGTGCTTTCACCGTTTATGCTGATCGGA 504

DB 306 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 247

QY 505 GATTTGACTGTGAAAAATCCCTGCTTTGTGTTTGTTCATATAATGAGATGAT 564

DB 246 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 187

QY 565 CTACCTTTTGTGCTTTGATGTTTGTGTTTGAACCTATGCTGTGCTGTATTAAT 624

DB 186 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 127

QY 625 TCACGTCATGTGAGATTGAGATTTT 653

DB 126 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 98

RESULT 4

ABV40063/C
ABV40063 standard; cDNA; 556 BP.

AC ABV40063;

DT 16-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 40054.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KM pharmacogenomic marker; gene; ss.

OS Homo sapiens.

PN WO200160860-A2.

PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.

PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Schlegel R, Endege WO, Monahan JE;

PD WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1, Page 8100; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

XX Query Match 9.9%; Score 68.6; DB 5; Length 556;
 XX Best Local Similarity 48.3%; Pred. No. 8.6e-06;
 XX Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

265 CTCGCTGTTACTGTAAGTTTCTGCTTGAAGCTCCGATCGCTCAGCGATGCA 324
DB CCCCCCTCCCCCTTTTATTTTATTTTATTTTCTCCCCCTCCCATTAATA 427
QY 325 TTCTGTCGATTTCTTTTCTGCTGGAATAATTCCTTAATGTCGATTTG 384
DB 426 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 367
QY 385 AAGTTTTCGATGAGGTTTCTTCTTATTTTATTTTATTTTATTTTATTTT 444
DB 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
QY 445 CTCGCTGTTACTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTG 504
DB 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
QY 505 GATTTGACGTGAAATTCCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTG 564
DB 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
QY 565 CTACCTTTGCTTGTGATTTGTTTGTGAGCCATAGCGTTGCTGTTTAACT 624
DB 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
QY 625 TCACGTTCAATGTGATTTGAGATTTT 653
DB 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98
  
```

RESULT 5

ABV42105/C

ID ABV42105 standard; cDNA; 556 BP.

XX ABV42105;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 42096.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-018319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Claim 1, Page 8441; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;

XX Query Match 9.9%; Score 68.6; DB 5; Length 556;
 XX Best Local Similarity 48.3%; Pred. No. 8.6e-06;
 XX Matches 188; Conservative 0; Mismatches 201; Indels 0; Gaps 0;

```

265 CTCGCTGTTACTGTAAGTTTCTGCTTGAAGCTCCGATCGCTCAGCGATGCA 324
DB 486 CCCCCCTCCCCCTTTTATTTTATTTTATTTTCTCCCCCTCCCATTAATA 427
QY 325 TTCTGTCGATTTCTTTTCTGCTGGAATAATTCCTTAATGTCGATTTG 384
DB 426 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 367
QY 385 AAGTTTTCGATGAGGTTTCTTCTTATTTTATTTTATTTTATTTTATTTT 444
DB 366 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 307
QY 445 CTCGCTGTTACTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTG 504
DB 306 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 247
QY 505 GATTTGACGTGAAATTCCTGTTTCTGTTTCTGTTTCTGTTTCTGTTTCTG 564
DB 246 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 187
QY 565 CTACCTTTGCTTGTGATTTGTTTGTGAGCCATAGCGTTGCTGTTTAACT 624
DB 186 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 127
QY 625 TCACGTTCAATGTGATTTGAGATTTT 653
DB 126 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 98
  
```

RESULT 6

ABV43601/C

ID ABV43601 standard; cDNA; 556 BP.

XX	ABV43601;
XX	16-SEP-2002 (first entry)
DT	Human prostate expression marker cDNA 43592.
DE	Human prostate expression marker CDNA 43592.
XX	Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW	pharmacogenomic marker; gene; ss.
XX	Homo sapiens.
OS	WO200160860-A2.
PN	23-AUG-2001.
XX	20-FEB-2001; 2001WO-US005171.
PD	17-FEB-2000; 2000US-0183319P.
XX	16-MAR-2000; 2000US-0189862P.
PR	25-MAY-2000; 2000US-0207454P.
PR	09-JUN-2000; 2000US-0211314P.
PR	18-JUL-2000; 2000US-0219007P.
PR	13-DEC-2000; 2000US-0255281P.
XX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA	Schlegel R, Endege WO, Monahan JE;
PI	WPI; 2001-662795/76.
DR	Novel isolated nucleic acid molecule associated with cancerous state of
XX	prostate cells and correlating with presence of prostate cancer, useful
XX	for detecting presence of prostate cancer, stage of prostate cancer.
PT	Claim 1; Page 8681-8682; 11750pp; English.
PS	The invention relates to an isolated nucleic acid molecule (1) comprising
XX	a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC	specification or its complement. (1) is useful for: (a) assessing whether
CC	a patient is afflicted with prostate cancer; (b) monitoring the
CC	progression of prostate cancer in a patient; (c) assessing the efficacy
CC	of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC	the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC	(e) selecting a composition for inhibiting prostate cancer in a patient;
CC	(f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC	determining whether prostate cancer has metastasized in a patient; (h)
CC	assessing the aggressiveness or indolence of prostate cancer in a patient
CC	; (1) is also useful as a pharmacodynamic or pharmacogenomic marker
CC	Sequence 556 BP; 389 A; 1 C; 51 G; 113 T; 0 U; 2 Other;
SQ	
Query Match	9.9%; Score 68.6; DB 5; Length 556;
Best Local Similarity	48.3%; Pred. NO. 8.6e-06;
Matches 188; Conservative	0; Mismatches 201; Indels 0; Gaps 0
Dy	265 CTCGCGCTTACTGTGAAGTTTTCGGCTTAGAGCCTCCGATGCGCTCACCGCATGCA 324
Dy	486 CNCCTCTCCCCCTTTTAATTATTTTTTTAATTTTCTCCCCCCCCTCCCAATAANA 427
Dy	325 TTCTGTCGTGATTTCTCTTTTCTTCGGTGAAAAATTCGCCCTAATGTTCTCGATTTGC 384
Dy	426 TTTTATTTTTTTTTTTTTTTTAAATAAAAACNCCCCCTTTTTTTTTTTTTTTTTT 367
Dy	385 AAGGTTTTTGCGTAAAGGTTACTTTTTTCCCTAATTTTAAGTTCTAGTAACGATA 444
Dy	366 TTTTATTTTTTTTTTTTTTTAATTTTCTTTTTTTTTTTAATTTTTTTTTTTTTTTT 307
Dy	445 CTGCGCTTACTGTTTTTGTTTCATTTTGTGTGCTTTCACCGTTAGTGGCTGATCGGA 504
Dy	306 TTTTATTTTTTTTTTTTTTTTTTTTTTTTCCCTTTTTTTTTTTTTTTTTTTTTTTTTT 247
Dy	505 GTAATTGACGTGAAAAACCTCGTTTTTGGTTTTGTTTCATATAAATCGAATGAT 564

[illegible]

Query Match 9.6%; Score 66.6; DB 4; Length 9539;
 Best Local Similarity 46.9%; Pred. No. 4.4e-05;
 Matches 207; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

251 TTTAACTACGCTCCGCTGCTTACCTGTAAGTTTTCGCTAGAGCCCTCGATCG 310
 202 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 261
 311 CCTACCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370
 262 GTTTTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
 371 TGTTCGATTTGCAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 430
 322 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 381
 431 CTGAGTAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 490
 382 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 441
 491 AGTCGCTGATGAGATTTTGAAGTGAAGAAATCCCTGCTGCTGCTGCTGCTGCTGCT 550
 442 ATTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 501
 551 TAAATGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610
 502 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 561
 611 GGCCTGTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670
 562 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 621
 671 TCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691
 622 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 642

RESULT 8
 ABRK28180
 ID ABRK28180 standard; DNA, 9539 BP.

AC ABRK28180;
 DT 23-APR-2002 (first entry)

DE DNA transcription associated complementary genomic DNA #27.

KM DNA transcription associated gene; peptide nucleic acid; PNA-oligomer;
 KM PNA; cytosine methylation state; SNP; retroviral infection; gene; ds;
 KM single nucleotide polymorphism; adenomine deaminase deficiency; cancer;
 KM viral infection; Sezary syndrome; haematological disorder; tuberculosis;
 KM immunological disorder; Werner syndrome; developmental disorder;
 KM psoriasis; Rieger's syndrome; neurological disorder; erythropoiesis;
 KM neurodegenerative disorder; Waardenburg syndrome; Niemann-Pick disease;
 KM myelodysplastic syndrome; myocardial infarction; hypertension; arthritis;
 KM angiogenesis; congenital heart disease; HDR syndrome; gene therapy;
 KM polyglutamine disorder; solid tumour.

OS Unidentified.

XX WO200192565-A2.

PN 06-DEC-2001.

PD 06-APR-2001; 2001WO-EP003973.

XX 06-APR-2000; 2000DE-01019058.

PR 07-APR-2000; 2000DE-01019173.

PR 30-JUN-2000; 2000DE-01032529.

PR 01-SEP-2000; 2000DE-01043826.

XX (EPIG-) EPIGENOMICS AG.

PI Olek A. Piepenbrock C. Berlin K;
 XX WPI, 2002-090046/12.

PT New nucleic acids or oligomers, useful for diagnosing or treating
 PT diseases associated with DNA transcription, e.g. immunological disorders,
 PT Werner syndrome, psoriasis, myocardial infarction, solid tumors or
 PT cancer.

PS Claim 1; SEQ ID NO 54; 32pp; English.

XX The invention relates to a nucleic acid, which comprises a segment of the
 CC chemically pretreated DNA of genes associated with DNA transcription from
 CC one of 346 sequences, and an oligomer, in particular an oligonucleotide
 CC or peptide nucleic acid (PNA)-oligomer that hybridises to or is identical
 CC to the chemically pretreated DNA of genes associated with DNA
 CC transcription. The set of oligomer probes are useful for detecting the
 CC cytosine methylation state and/or single nucleotide polymorphisms (SNPs)
 CC in a chemically pretreated genomic DNA. The nucleic acids are useful for
 CC diagnosing or treating diseases associated with DNA transcription
 CC (particularly with the methylation status), e.g. adenomine deaminase
 CC deficiency, viral infection, retroviral infection, Sezary syndrome,
 CC haematological disorders, immunological disorders, Werner syndrome,
 CC tuberculosis, developmental disorders, psoriasis, Rieger's syndrome,
 CC neurodegenerative disorders, neurodegenerative disorders, Waardenburg
 CC syndrome, Niemann-Pick disease, myelodysplastic syndrome, myocardial
 CC infection, hypertension, angiogenesis, erythropoiesis, congenital heart
 CC disease, HDR syndrome, arthritis, polyglutamine disorders, solid tumours
 CC or cancer. Sequences ABRK28127-ABRK28472 represent DNA transcription
 CC associated genomic DNA molecules of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification but
 CC was obtained in electronic format directly from the European Patent
 CC Office

SO Sequence 9539 BP; 2078 A; 121 C; 2281 G; 5059 T; 0 U; 0 Other;

Query Match 9.6%; Score 66.6; DB 6; Length 9539;
 Best Local Similarity 46.9%; Pred. No. 4.4e-05;
 Matches 207; Conservative 0; Mismatches 234; Indels 0; Gaps 0;

251 TTTAACTACGCTCCGCTGCTTACCTGTAAGTTTTCGCTAGAGCCCTCGATCG 310
 202 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 261
 311 CCTACCGCATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 370
 262 GTTTTATTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 321
 371 TGTTCGATTTGCAAGTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 430
 322 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 381
 431 CTGAGTAAAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 490
 382 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 441
 491 AGTCGCTGATGAGATTTTGAAGTGAAGAAATCCCTGCTGCTGCTGCTGCTGCTGCT 550
 442 ATTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 501
 551 TAAATGAGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 610
 502 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 561
 611 GGCCTGTTAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 670
 562 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 621
 671 TCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 691
 622 TTTTATTTTGTGATTTTTCGTTTATTTTGTGCTTTTATTTTATTTTATTTTTC 642

```
RESULT 9
AA33181/c
ID AAX33181 standard; DNA; 6644 BP.
XX
XX AAX33181;
XX
XX 25-JUN-1999 (first entry)
XX
XX Base sequence of the plasmid prx-ires-bsr.
DE
XX
XX CompoX virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
XX bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer;
XX autoimmune disease; graft rejection reaction; inflammation;
XX inflammatory disease; ss.
XX
XX Synthetic.
OS CompoX virus.
XX
XX WO9913073-A2.
XX
XX 18-MAR-1999.
XX
XX 07-SEP-1998; 98WO-IP004010.
XX
XX 08-SEP-1997; 97JP-00259235.
XX
XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX
XX Hamada H;
XX
XX WPI; 1999-243728/20.
XX
XX New apoptosis-resistant virus-sensitive cell.
XX
XX Example 1; Page 38-41; 51pp; English.
XX
XX The present invention describes an apoptosis-resistant virus-sensitive
XX cell line into which an apoptosis resistance gene has been introduced.
XX The recombinant viruses generated are capable of expressing apoptosis-
XX associated genes. These can then be used in a variety of diseases for
XX which the induction of apoptosis is therapeutic, or where the
XX inhibition of harmful apoptosis is therapeutic. The recombinant viruses
XX are useful as vectors for gene therapy which can be applied to cancer
XX therapy for destroying cancer cells selectively, the treatment of
XX autoimmune diseases and graft rejection reaction, and apoptosis induction
XX therapy for inflammatory cells in inflammatory diseases. Prior arts have
XX encountered the problem where if an adenovirus vector capable of
XX expressing an apoptosis-associated gene is introduced into animal cells,
XX the cells producing the virus will be destroyed because the period of
XX time required to induce cell death by apoptosis is shorter than that
XX required to replicate and produce the virus, resulting in failure to
XX obtain a recombinant virus having the integrated apoptosis-associated
XX gene. In this invention an apoptosis-resistant 293 cell line (having an
XX apoptosis resistant gene introduced) is established and overcomes the
XX problem. The present sequence represents the base sequence of the plasmid
XX prx-ires-bar, which contains the compoX virus bsr gene, and is used in an
XX example from the present invention
XX
XX Sequence 6644 BP; 2166 A; 1573 C; 1424 G; 1481 T; 0 U; 0 Other;
XX
XX Query Match 9.5%; Score 66; DB 2; Length 6644;
XX Best Local Similarity 45.4%; Pred. No. 5.6e-05;
XX Matches 227; Conservative 0; Mismatches 285; Indels 0; Gaps 0;
XX
XX 157 TTTTAGAGCTTAATAAGCTTCTCAATTTGTCTCTTCGTCAGTTTATTTCTTCCTCC 216
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4235 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 4176
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 217 GGAGTGCTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 276
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 4175 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 4116
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
XX 277 CTCTGTAAGTTTCTGCTTAGAGCTCGATGCGCTCAACCGATGCAATTTGTGCTCGA 336
```

```
Db 4115 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 4056
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 337 TTTCTCTTTCTTTCTGCGGAAAAATGCCCTAATGTTCTGATTTGCAAGTTTGTG 396
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 4055 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3996
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 397 CTATGGTTACTTTTTCCTTAATTTTATAGTCTTAGATACGATACGATGAC 456
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3995 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3936
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 457 TCTTTTGTTCATTTTGTGTCCTTACCGTTTAGTGTGAGAGATTTGACGT 516
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3935 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3876
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 517 GAAAAATCCTTCGTTTGTGTTTGTGTTTGTTCATATTAACGATGATGACCTTTGTG 576
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3875 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3816
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 577 CTATGATGTTGTTTGTGAGCCCTAAGCGTTGCGCTTGTATTAACCTACGTTTCANGT 636
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3815 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3756
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
QY 637 GTGATTTTGAGATTTTGTGATGACGTCGTGCTTTCTTTGCT 678
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 3755 TTTTCTTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTTCTTT 3714
XX ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 10
AA33182/c
ID AAX33182 standard; DNA; 7372 BP.
XX
XX AAX33182;
XX
XX 25-JUN-1999 (first entry)
XX
XX Base sequence of the plasmid prx-bcl-x1-bsr.
DE
XX
XX CompoX virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
XX bcl-2; bcl-x1; FLIP; survivin; IAP; ILP; adenovirus; cancer;
XX autoimmune disease; graft rejection reaction; inflammation;
XX inflammatory disease; ss.
XX
XX Synthetic.
OS Homo sapiens.
XX
XX WO9913073-A2.
XX
XX 18-MAR-1999.
XX
XX 07-SEP-1998; 98WO-IP004010.
XX
XX 08-SEP-1997; 97JP-00259235.
XX
XX (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX
XX Hamada H;
XX
XX WPI; 1999-243728/20.
XX
XX New apoptosis-resistant virus-sensitive cell.
XX
XX Example 2; Page 41-45; 51pp; English.
XX
XX The present invention describes an apoptosis-resistant virus-sensitive
XX cell line into which an apoptosis resistance gene has been introduced.
XX The recombinant viruses generated are capable of expressing apoptosis-
XX associated genes. These can then be used in a variety of diseases for
XX which the induction of apoptosis is therapeutic, or where the
XX inhibition of harmful apoptosis is therapeutic. The recombinant viruses
XX are useful as vectors for gene therapy which can be applied to cancer
XX therapy for destroying cancer cells selectively, the treatment of
XX autoimmune diseases and graft rejection reaction, and apoptosis induction
```

therapy for inflammatory cells in inflammatory diseases. Prior arts have encountered the problem where if an adenovirus vector capable of expressing an apoptosis-associated gene is introduced into animal cells, the cells producing the virus will be destroyed because the period of time required to induce cell death by apoptosis is shorter than that required to replicate and produce the virus, resulting in failure to obtain a recombinant virus having the integrated apoptosis-associated gene. In this invention an apoptosis-resistant 293 cell line (having an apoptosis resistant gene introduced) is established and overcomes the problem. The present sequence represents the base sequence of the plasmid pRxBcl-x1-ber, which contains the human Bcl-x1 gene, and is used in an example from the present invention

Sequence 7372 BP; 2353 A; 1749 C; 1649 G; 1621 T; 0 U; 0 Other;

Query Match 9.5%; Score 66; DB 2; Length 7372;
Best Local Similarity 45.4%; Pred. No. 5.7e-05;
Matches 237; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

157 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 216
|||||
4963 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4904
217 GGAGTCTGACCTACACTCTGCTCGGCGCTTAAAGCTTCCGCTCTTA 276
|||||
4903 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4844
277 CTCGTAAAGCTTCTGCTGCTTAAAGCTTCCGCTTAAAGCTTCTGCTCTGA 336
|||||
4843 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4784
337 TTTCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 396
|||||
4783 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4724
397 CTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 456
|||||
4723 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4664
457 TGTATGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 516
|||||
4663 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4604
517 GAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 576
4603 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4544
577 TGTATGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 636
|||||
4543 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4484
637 GTGAGTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 678
4483 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4442

RESULT 11
AA33180/c
ID AA33180 standard; DNA; 7797 BP.

AA33180;

25-JUN-1999 (first entry)

Compox virus bsr full length gene sequence.

Compox virus; bsr: viral vector; expression; apoptosis; resistance; crna;

bcl-2; bcl-x1; FltP; survival; IAP; ILP; adenovirus; cancer;

autoimmune disease; graft rejection reaction; inflammation;

inflammatory disease; ss.

Compox virus.

XX

PN WO913073-A2.
XX 18-MAR-1999.
PD 07-SEP-1998; 98WO-JP004010.
PF 08-SEP-1997; 97JP-00259235.
XX (RPRG-) RPR GENCBLL ASIA PACIFIC INC.
XX Hamada H;
XX WPI; 1999-243728/20.
XX New apoptosis-resistant virus-sensitive cell.
XX Example 1; Page 34-38; 51pp; English.
XX The present invention describes an apoptosis-resistant virus-sensitive
CC cell line into which an apoptosis resistance gene has been introduced.
CC The recombinant viruses generated are capable of expressing apoptosis-
CC associated genes. These can then be used in a variety of diseases for
CC which the induction of apoptosis by gene transfer, or where the
CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
CC are useful as vectors for gene therapy which can be applied to cancer
CC therapy for destroying cancer cells selectively, the treatment of
CC autoimmune diseases and graft rejection reaction, and apoptosis induction
CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
CC encountered the problem where if an adenovirus vector capable of
CC expressing an apoptosis-associated gene is introduced into animal cells,
CC the cells producing the virus will be destroyed because the period of
CC time required to induce cell death by apoptosis is shorter than that
CC required to replicate and produce the virus, resulting in failure to
CC obtain a recombinant virus having the integrated apoptosis-associated
CC gene. In this invention an apoptosis-resistant 293 cell line (having an
CC apoptosis resistant gene introduced) is established and overcomes the
CC problem. The present sequence represents the compox virus bsr gene which
CC is used in an example from the present invention

Sequence 7797 BP; 2542 A; 1760 C; 1656 G; 1639 T; 0 U; 0 Other;

Query Match 9.5%; Score 66; DB 2; Length 7797;
Best Local Similarity 45.4%; Pred. No. 5.8e-05;
Matches 237; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

157 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 216
|||||
5388 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 5329
217 GGAGTCTGACCTACACTCTGCTCGGCGCTTAAAGCTTCTGCTCTGA 276
|||||
5328 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 5269
277 CTCGTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 336
|||||
5268 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 5209
337 TTTCTCTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 396
|||||
5208 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 5149
5148 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 5089
397 CTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 456
|||||
5148 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 5029
457 TGTATGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 516
|||||
5088 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 5029
5088 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 5029
517 GAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 576
|||||
5028 TTTTAAAGCTTAAAGCTTCCATTTGCTCTCTGCTGAGTTATTTCTTCCCTC 4969

```

Oy 577 CTTGATGTTGTTTGTGACCTATGCGTGTGCTGTATTAACCTCAGTCATGT 636
Db 4968 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4909
Oy 637 GTGGATTTGAGATTTTGTGAGTCTGCGGTTCTTTGCT 678
Db 4908 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 4867

RESULT 12
AAK3184/c
ID AAK3184 standard; DNA; 7996 BP.
XX
AC AAK3184;
XX
DT 25-JUN-1999 (first entry)
XX
DE Base sequence of the plasmid pRX-Bcl 2-1-hCD 25.
XX
KM CompoX virus; bsr; viral vector; expression; apoptosis; resistance; crmA;
KM bcl-2; bcl-x1; FLIP; survivin; IAP; IAP; adenovirus; cancer;
KM autoimmune disease; graft rejection reaction; inflammation;
KM inflammatory disease; ss.
XX
OS Synthetic.
XX
OS Homo sapiens.
XX
PN MO913073-A2.
XX
PD 18-MAR-1999.
XX
PF 07-SEP-1998; 98WO-JP004010.
XX
PR 08-SEP-1997; 97JP-00259235.
XX
PA (RPRG-) RPR GENCELL ASIA PACIFIC INC.
XX
PI Hamada H;
XX
DR WPI; 1999-243728/20.
XX
XX
XX New apoptosis-resistant virus-sensitive cell.
XX
XX Example 3; Page 46-49; 51pp; English.
XX
CC The present invention describes an apoptosis-resistant virus-sensitive
CC cell line into which an apoptosis resistance gene has been introduced.
CC The recombinant viruses generated are capable of expressing apoptosis-
CC associated genes. These can then be used in a variety of diseases for
CC which the induction of apoptosis by gene transfer, or where the
CC inhibition of harmful apoptosis, is therapeutic. The recombinant viruses
CC are useful as vectors for gene therapy which can be applied to cancer
CC therapy for destroying cancer cells selectively, the treatment of
CC autoimmune diseases and graft rejection reaction, and apoptosis induction
CC therapy for inflammatory cells in inflammatory diseases. Prior arts have
CC encountered the problem where if an adenovirus vector capable of
CC expressing an apoptosis-associated gene is introduced into animal cells,
CC the cells producing the virus will be destroyed because the period of
CC time required to induce cell death by apoptosis is shorter than that
CC required to replicate and produce the virus, resulting in failure to
CC obtain a recombinant virus having the integrated apoptosis-associated
CC gene. In this invention an apoptosis-resistant 293 cell line (having an
CC apoptosis resistant gene introduced) is established and overcomes the
CC problem. The present sequence represents the base sequence of the plasmid
CC pRX-Bcl 2-1-hCD 25, which contains the human Bcl-2 gene, and is used in
CC an example from the present invention
XX
SQ Sequence 7996 BP; 2463 A; 2015 C; 1829 G; 1689 T; 0 U; 0 Other;

```

```

Query Match 9.5%; Score 66; DB 2; Length 7996;
Best Local Similarity 45.4%; Freq. No. 5.8e-05;
Matches 237; Conservative 0; Mismatches 285; Indels 0; Gaps 0;

```

```

Oy 157 TTTAGAGCTTAATAAGCTTCCTCATTTGCTCTCTGTCAGTTATTTCTCTCC 216
Db 5587 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5528
Oy 217 GGAGCTCAGCAGTCTACTCTCAGCTCCGGGCTTTAACTAGCTCTCCGCTTAA 276
Db 5527 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5468
Oy 277 CTCTGTAAGTTTCTGCGCTTAGAGCGCTCGAGTCGCTCAGCGATTCGTGCTGA 336
Db 5467 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5408
Oy 337 TTTCTCTTTTCTGCGCTGGAATAATGCTCATGTCGATTTGGAAGTTTGTG 396
Db 5407 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5348
Oy 397 CTATGGGTTACTTTTCCCTAATTTAATGTTCTTAGCAATACGATACCTGCGCTTAC 456
Db 5347 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5288
Oy 457 TGTGTGTGTGTCATTTGTTGCTGCTTACCGCTTGTAGCGCTGATGCGAGTATTTGACTGT 516
Db 5287 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5228
Oy 517 GAAAAATCCTGCTTTTGTGTTTGTGTTTCATATTAATCGAATTGATCTACCTTTGTG 576
Db 5227 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5168
Oy 577 CTTGATGTTGTTTGTGAGCCTATGCGTGTGCTGTATTAACCTCAGTCATGT 636
Db 5167 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5108
Oy 637 GTGGATTTGAGATTTTGTGAGTCTGCGGTTCTTTGCT 678
Db 5107 TTTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 5066

```

```

RESULT 13
ABL33958
ID ABL33958 standard; DNA; 14006 BP.
XX
AC ABL33958;
XX
XX
XX 26-MAR-2002 (first entry)
XX
DE Human immune system associated gene SEQ ID NO: 1931.
XX
XX Human; immune system disease; cytosine methylation; antiaesthetic;
XX antiarteriosclerotic; antianaemic; cytosarctic; nocrotropic;
XX neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
XX antineumatic; antirathritic; antidiabetic; antipsoriatic;
XX antineuroinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
XX acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
XX neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease; gene;
XX ds.
XX
XX Homo sapiens.
XX
XX WO200200928-A2.
XX
XX 03-JAN-2002.
XX
XX 02-JUL-2001; 2001WO-BP007537.
XX
XX 30-JUN-2000; 2000DE-01032529.
XX
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIC-) EPICENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
XX

```


RESULT 15

AB036996 standard; DNA; 556 BP.

AB036996;

12-JUL-2002 (first entry)

Oligonucleotide for detecting cytosine methylation SEQ ID NO 23587.

Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;

drug; side effect; cancer; central nervous system; cardiovascular;

gastrointestinal; respiratory system; single nucleotide polymorphism;

SNP; cell differentiation; ds

Homo sapiens.

WO200218632-A2.

01-SEP-2001; 2001WO-EP010074.

01-SEP-2000; 2000DE-01043826.

05-SEP-2000; 2000DE-01044543.

(EPIC-) EPIGENOMICS AG.

Olek A, Piepenbrock C, Berlin K, Gueitig D;

MPI; 2002-371829/40.

Determining the degree of cytosine methylation in genomic DNA, useful for

diagnosis and prognosis, comprises selective hybridization of amplicons

from chemically treated DNA.

Claim 12; 56pp + Sequence Listing; 56pp; German.

This invention describes a novel method for determining the degree of

methylation of a particular cytosine in a motif 5'-CpG-3', present in a

genomic sample of DNA. The sample is treated chemically to convert

cytosine (C) but not methylated C, to uracil, then part of the genomic

DNA that contains the target C is amplified to form a labeled amplicon.

The amplicon is hybridized to two classes, each with at least one member,

of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the

degree of hybridization to both classes is determined from the label on

the amplicon. From the ratio of labels hybridized to the two classes of

oligomers, the degree of methylation is calculated. The method is used:

(i) for diagnosis and/or prognosis of side effects of therapeutic drugs

and of a wide range of diseases, e.g. cancer, disorders of the central

nervous, cardiovascular, gastrointestinal and respiratory systems etc.,

particularly by detecting mutations or single-nucleotide polymorphisms

(SNP/s); and (ii) for differentiation of cell or tissue types and for

investigating cell differentiation. The method allows the methylation

status of many C residues to be determined simultaneously. AB013410-

AB054121 represent genomic DNA sequences used to illustrate the method

for determining the degree of cytosine methylation described in the

disclosure of the invention

Sequence 556 BP; 12 A; 13 C; 14 G; 505 T; 0 U; 12 Other;

Query Match 9.5%; Score 65.8; DB 6; Length 556;

Best Local Similarity 49.3%; Pred. No. 3.6e-05;

Matches 172; Conservative 0; Mismatches 177; Indels 0; Gaps 0;

325 TTCTGAGCTGATTTCTCTTTTCTGCGTGAATAATGCCCTATGTTCTGATTCG 384

96 TTTCGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTATTTTTTTTTTTT 155

385 AAGGTTTTTGCTAGGCTTACTTTTTCCCTATATTTATATTTCTTAGGTAACGATA 444

Db 156 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTACGTTTTTTTTTCGTTTTTTTTTTTTT 215

QY 445 CCGGCTTTACGTTTTTTTTCATTTTGTGTCGCTTCACCGTTTAGTCGATTCGA 504

Db 216 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTT 275

QY 505 GTATTTCACCTGAAAAATCCTTCGTTTTTTTGTTCATATTAATCGATTCAT 564

Db 276 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTT 335

QY 565 CTACCTTTTGTGCTTGTGATTTGTTTTTGAACCTATGCGTTGCTGTATTA 624

Db 336 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTT 395

QY 625 TCACGTTATGTCGATTTGATTTGATTTGATTTGATTTGATTTGATTTGATTT 673

Db 396 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTTTTTTTTTTT 444

Search completed: April 8, 2004, 09:18:38
Job time : 184.901 secs

This Page Blank (uspto)

TITLE Direct Submission
JOURNAL Submitted (01-JUL-2003) Weisshaar B., Max-Planck-Institut fuer Zuechtungsforchung, Carl-von-Linne-Weg 10, Koeln, 50823, Germany
COMMENT This sequence is recovered from the left border of the T-DNA. It indicates an insertion within the locus defined by clone M0M1. The sequences are generated at the MPI for Plant Breeding Research in the context of the GABI-Kat project. GABI-Kat is part of the German Plant Genomics program designated 'GABI'. Information on line availability can be found at:
<http://www.mpiz-koeln.mpg.de/GABI-Kat/>.

	Query Match	Similarity	52.5%;	Score 364.6;	DB 29;	Length 395;	
	Best Local	Matches	376;	Conservative	0;	Mismatches 19;	Indels 0;
QY	114	ACTGAATGAAAATTTGGAGTCCGAATTCGGAATAACGAGCCGTTTAAAGCTTAATAG	173				
Db	1	ACTGATGAAAATTTGGAGTCGGAATTCGAAAACGAGCCGTTTAAAGCTTAATAG	60				
QY	174	CTTCCCATTTGCTCTCTCTGTCGATTAATTTCTTCTCCGAGTCCTGACTCACTA	233				
Db	61	CTTCCCATTTGCTCTCTCTGTCGATTAATTTCTTCTCCGAGTCCTGACTCACTA	120				
QY	234	CTCTCACTCTCCGCGCTTTAAACTTAAGTTCCTCGCTTTACTCTGTAAGTTTCTCG	293				
Db	121	CTCTCACTCTCCGCGCTTTAAACTTAAGTTCCTCGCTTTACTCTGTAAGTTTCTCG	180				
QY	294	CTTGAAGCTCCGATTCGCTCACCGCATGGAATCTGTCGATTCGATTCCTTTTCTTCGC	353				
Db	181	CTTGAAGCTCCGATTCGCTCACCGCATGGAATCTGTCGATTCGATTCCTTTTCTTCGC	240				
QY	354	TGGAATAATGACCCTAATGTTCTCGATTTGAAAGTTTGTGCTAGGTTTACTTTT	413				
Db	241	TGGAATAATGACCCTAATGTTCTCGATTTGAAAGTTTGTGCTAGGTTTACTTTT	300				
QY	414	CCCTAATATTTAATGTTCTTAAGTGAAGTAACCTGCGCTTACGTTTTGTTCATTTG	473				
Db	301	CCCTAATATTTAATGTTCTTAAGTGAAGTAACCTGCGCATTAAGTTTGTTCATTTG	360				
QY	474	TTTGCTTTCACGTTTATAGTCGTATCGAGATTT	508				
Db	361	TTTGCTTTCACGTTTATAGTCGTATCGAGATTT	395				

RESULT 2	303 bp	DNA	linear	GS5-02-UTL-2001
BX547724				
LOCUS				
DEFINITION	Arabidopsis thaliana T-DNA flanking sequence GK-547E01-020586,			
	genomic survey sequence.			
ACCESSION	BX547724.1	GI:32440533		
VERSION				
KEYWORDS				
SOURCE	GS5.			
ORGANISM	Arabidopsis thaliana (thale cress)			
	Arabidopsis thaliana			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			

REFERENCE AUTHORS TITLE	JOURNAL REFERENCE AUTHORS TITLE
1 Strizhov, N., Li, Y., Rosso, M., Viehoveer, P., Dekker, K., Saedler, H. and Weishaar, B. A pipeline for automated high-throughput generation of FSTS (flanking sequence tags) from Arabidopsis thaliana T-DNA transformed lines unpublished	2 Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weishaar, B. Rosso, M., Strizhov, N., Li, Y., Reiss, B., Dekker, K. and Weishaar, B. A new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat) and a new Arabidopsis thaliana T-DNA mutagenised population (GABI-Kat)

FEATURES
SOURCE

Query Match	39.9%	Score 277.4	DB 29	Length 303
Best Local Similarity	94.7%	Pred. No. 55-35		
Matches 287		Mismatches 0	Indels 0	Gaps 0
		Conservative		

QY	415	CCT	417
Db	301	CCT	303

RESULT 3
BH854770 209 bp DNA linear GSS 08-JUL-2002
LOCUS BH854770
DEFINITION SALK_088648.19.55.x Arabidopsis thaliana TDNA insertion line
Arabidopsis thaliana genomic clone SALK_088648.19.55.x, genomic survey sequence.
ACCESSION BH854770
VERSION BH854770.1 GI:21704360
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 209)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadinb,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmermann,J. and Ecker,J.R.
A Sequence-indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
JOURNAL Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.
FEATURES
source Location/Qualifiers
1..209
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_088648.19.55.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
ORIGIN
Query Match 22.7%; Score 157.8; DB 28; Length 209;
Best Local Similarity 84.7%; Pred. No. 2.6e-16;
Matches 177; Conservative 0; Mismatches 32; Indels 0; Gaps 0;
QY 322 GCATTGCTGCTGATTTCTTTTCTTGCTGGAATAATGCGCTAATGTTCTGATT 381
DB 1 GCATTGCTGCTGATTTCTTTTCTTGCTGGAATAATGCGCTAATGTTCTGATT 60
QY 382 TCGAAGGTTTGTGCTATGCGTTACTTTTCCATATATCTATAGTGAGGCGG 441
DB 61 TCGAAGGTTTGTGCTATGCGTTACTTTTCCATATATCTATAGTGAGGCGG 120
QY 442 ATACCTGCTTACTGTTTGTTCATTTTGTGCTTTCACCGTTAGTGCCTGATC 501
DB 121 ATGCGCGCTTATAGGCTATGAAGTTCGGGCGCTTTCCCGGTTAGCGCGTATC 180
QY 502 GGAATTTTGAATGCTGGAATAATCTTCTGT 530
DB 181 GGAATTTTGAATGCTGGAATAATCTTCTGT 209
RESULT 4
BH847287 144 bp DNA linear GSS 13-JUN-2002
LOCUS BH847287
DEFINITION SALK_050870.18.25.x Arabidopsis thaliana TDNA insertion line

Arabidopsis thaliana genomic clone SALK_050870.18.25.x, genomic survey sequence.
ACCESSION BH847287
VERSION BH847287.1 GI:21418158
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 144)
Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R., Gadinb,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L., Shinn,P., Zimmermann,J. and Ecker,J.R.
A Sequence-indexed Library of Insertion Mutations in the Arabidopsis Genome
Unpublished (2001)
JOURNAL Contact: Joseph R. Ecker
Salk Institute Genomic Analysis Laboratory (SIGAL)
The Salk Institute for Biological Studies
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
Tel: 858 453 4100 x1752
Fax: 858 558 6379
Email: ecker@salk.edu
This is single pass sequence recovered from the left border of TDNA.
FEATURES
source Location/Qualifiers
1..144
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/strain="Columbia 0"
/db_xref="taxon:3702"
/clone="SALK_050870.18.25.x"
/clone_lib="Arabidopsis thaliana TDNA insertion lines"
/note="PCR was performed on Arabidopsis thaliana lines each of which contains one or more TDNA insertion elements. The resultant fragment for each line was directly sequenced to determine the genomic sequence at the site of insertion. Details of the protocols used can be found at http://signal.salk.edu/tdna_protocols.html"
ORIGIN
Query Match 15.7%; Score 108.8; DB 28; Length 144;
Best Local Similarity 84.7%; Pred. No. 3.6e-08;
Matches 122; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
QY 372 GTTCGATTTGGAAGTTTGTGCTATGCTTCTTTTCCATATTTTATAGTTC 431
DB 1 GTTCGATTTGGAAGTTTGTGCTATGCTTCTTTTCCATATTTTATAGTTC 60
QY 432 TTAGTAAGATACCTGCTTACTGTTTGTTCATTTGTGTGCTTACCGTTTA 491
DB 61 TTAGTAAGATACCTGCTTACTGTTTGTGTGCTTGTGCGATTTGCGCGCTGA 120
QY 492 GTGCGATCGAGTATTGACTG 515
DB 121 GTGCGATCGAGTATTGACTG 144
RESULT 5
A0964581 232 bp DNA linear GSS 28-JAN-2000
LOCUS A0964581
DEFINITION LERGX20TR LERX Arabidopsis thaliana genomic clone LERGX20, genomic survey sequence.
ACCESSION A0964581
VERSION A0964581.1 GI:6792282
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.

ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 922)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..922
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR14D09"
/clone_1ib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 12.7%; Score 88.4; DB 29; Length 922;
Best Local Similarity 29.4%; Pred. No. 2.3e-05;
Matches 200; Conservative 165; Mismatches 309; Indels 6; Gaps 2;

QY 2 ACAACAACAAATTAAGCAGTCACTGTTGAGCAATTAATTAAGTGGAAACAT 61
DB 239 AAAAAACWCTAACTAATAATTAACAAATAMCAACAAATTAATTAACAAAAACAA 298
QY 62 TAAGTTAAGGAAAGAAAAAGTAAAGTAAAGTAAAGTAAAGTAAAGTAAAGT 121
DB 299 CAAAAAACCAAAAAAANAAAAAANAAAAAANAAAAAACAACAAACAAAAACAA 358
QY 122 AAAATTGAGTCAGATCGAAAAACGAGCGCTTTAGAGCTTAATTAAGCTCTCA 181
DB 359 AAAA-----AACAAAAAATAAABACCCCAAAAAAACAACATCAAAATCTTT 413
QY 182 TTGTCTCTTCTGTCAGTTAATTTCTCTCCGAGTCGAGTCACTCACTCACT 241
DB 414 CTTCGCTTCCCTCTCTCTGTCCTGTCCTGTCCTCTCTCTCTCTCTCTCTCTCT 473
QY 242 CTCGCGCGCTTAACCTAAGTCTCTCGCTTAACTAGTCTAG-TTTTCTGCTAGAG 300
DB 474 CTGGGCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCT 533
QY 301 CCTCCAGTCGCTCACCGCAGTCACTCTGTCGATTTCTCTTTTCTCTGCGGAAAA 360
DB 534 GGTGKGSTKTTTCTTGTGKGTGKGTGKTTKTTKTTSTGGKKKTTTGTGKGTG 593
QY 361 ATTGCCCTAATGTTCTGATTTGAAAGTTTGTGCTAGTGGTTACTTTTCTCCATA 420
DB 594 TGGGKGGKGTGTTTGTGKGTGKGTGKTTKTTKTTKTTKTTKTTKTTKTTKTTG 653
QY 421 TTTTATAGTTCTAAGTAACGATACCGCTTACGTTTCTTGTCTATTTGTTGCT 480
DB 654 TTGGGKKTGTTTGTGKGTGKGTGKGTGKGTGKGTGKGTGKGTGKGTGKGTG 713
QY 481 TTCAACGTTTGTGCTGATCGAGATTTGACGTGAAAAATCTTCTGTTTGTGTT 540
DB 162 GAGCTTAATAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 221

DB 714 TTKTTTXXXXXXTTTXXXTTXXXTTXXXTTXXXTTXXXTTXXXTTXXXTTXXXT 773
QY 541 TTGTTTCATATAATGAGTATGATCACTTTGTCCTTTGATGTTGTTTTCAGCCT 600
DB 774 KKK 833
QY 601 ATGCGTTGTCCTTTTAACTTCACTGATGATGATGATGATGATGATGATGATG 660
DB 834 GTKKGTGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGGKGG 893
QY 661 ACTGTGCGTTTCTTTGCTG 680
DB 894 KKKKKKKKKKKKKGGGGGGG 913

RESULT 8
CNS005TE 997 bp DNA linear GSS 03-JUN-1999
LOCUS Drosophila melanogaster genome survey sequence TET3 end of BAC #
DEFINITION BACR12K22 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
ACCESSION AL060767.1 GI:4943573
VERSION AL060767.1
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 997)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mammosser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES
source
1..997
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR12K22"
/clone_1ib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 12.7%; Score 88; DB 29; Length 997;
Best Local Similarity 20.8%; Pred. No. 2.5e-05;
Matches 115; Conservative 160; Mismatches 278; Indels 0; Gaps 0;

QY 42 AAATTATAGTGGGAACAATTAAAGTGAAGAAAAAGAAAAAGTACAAAAATGA 101
DB 335 ATATNAANNNNAN 394
QY 102 AAACAAATCAACGAATGAATTTGAGATCGAATCGAAAAACGAGCGCTTTTA 161
DB 395 AAAAAAANNNATATGNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 454
QY 162 GAGCTTAATAGCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 221

[illegible]

RESULT	9
LOCUS	BX416727
DEFINITION	BX416727 Homo sapiens NEUROBLASTOMA 712 bp mRNA linear clone
ACCESSION	BX416727
VERSION	CSD0DA011Y114-5-PRIME, mRNA sequence.
KEYWORDS	BX416727.1 GI:30765629
SOURCE	EST.
ORGANISM	Homo sapiens (human)
REFERENCE	Homo sapiens; Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Euteria; Primates; Catarrhini; Hominidae; Homo. 1. (bases 1 to 712) Li, W.B., Gruber, C., Jesssee, J. and Polayes, D. Full-length cDNA libraries and normalization Unpublished (2001)
AUTHORS	Contact: Genoscope
TITLE	Genoscope - Centre National de Sequencage
JOURNAL	Bp 191 91006 EVRY cedex - France
COMMENT	Email: seque@genoscope.cns.fr Web : www.genoscope.cns.fr Library was constructed by Life Technologies, a division of Invitrogen. Contact : Peng liang Email : filiang@lifetech.com URL : http://fulllength.invitrogen.com/ Invitrogen Corporation 1600 Paradey Avenue Genoscope sequence ID : CSD0DA011BE07Qp1.
FEATURES	
source	1..712 Location/Qualifiers

```

/db_xref="taxon:9606"
/clone="CS0DA011Y114"
/tissue_type="NEUROBLASTOMA"
/clone_lib="Homo sapiens NEUROBLASTOMA"
/notes="Vector: pCMVSPORT 6; 1st strand cDNA was primed
with a NotI-oligo(dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector.
Library was not normalized."

```

Matches	54;	Conservative	185;	Mismatches	129;	Indels	0;	Gaps	0;
QY	324	ATTCTGCTCGATTTCTTTTCTTCGCTGAAAAATGCCCTAATGTTCTGATTC	383						
	:: ::	:: :: :: :: :: :: :: :: ::							
Db	330	MMMT	389						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
QY	384	GAAGTTTGTGTAGGTTACTTTTTCCTATATTATAGTCTTAGTAAGAT	443						
	:: ::	:: :: :: :: :: :: :: :: :: :: ::							
Db	390	TT	449						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
QY	444	ACGCGCTACAGTTTGTTCATTTGTGTGCTTACGCTTGAGCTGATCG	503						
	:: ::	:: :: :: :: :: :: :: :: :: ::							
Db	450	KTTTKTKKKKKKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTKTK	509						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
QY	504	AGATTTGACGTGAAAAATCCTTCGTTTTTGTGTTTTGTTTCATATAATCGATTGA	563						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
Db	510	KKK	569						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
QY	564	TCTACATTTGTGTGATGTTGTTGTTTGTGACCTATGCGTTGTTGATTAAC	623						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
Db	570	KKK	629						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
QY	624	TTACAGTTACGTGATTTTGTGATTTTGTAGATGCTGAGATTCTTTGTGAGTA	683						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
Db	630	KKK	689						
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
QY	684	TAGGTTGT 691							
	:: ::	:: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: :: ::							
Db	690	KKKKKKKK 697							

RESULT	10
LOCUS	AQ964580/c
DEFINITION	Arabidopsis thaliana genomic clone LERGX0, genomic survey sequence.
ACCESSION	AQ964580
VERSION	AQ964580.1 GI:6792281
KEYWORDS	GSS.
SOURCE	Arabidopsis thaliana (thale cress)
ORGANISM	Arabidopsis thaliana Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi- s. (bases 1 to 613) Buell C.R., Lin X., Pai G., Barnstead M., Bowman C., Uteerbach T., Feldblum T., Liang F., Creasy T. and Fraser C.M. Genomic survey sequencing of Landsberg erecta ecotype of Arabidopsis thaliana and identification of sequence-based polymorphisms unpublished (2000) Contact: Xiaoying Lin The Institute for Genomic Research 9712 Medical Center Dr., Rockville, MD 20850, USA Tel.: 301 838 0200 Fax: 301 838 0208 Email: atc@tigr.org For additional information, see http://www.tigr.org/tdb/at/at.html Seq primer: TF Class: Shotgun.

FEATURES	source	Location/Qualifiers
ORIGIN		1..613
		/organism="Arabidopsis thaliana"
		/mol_type="genomic DNA"
		/strain="LANDSBERG_ERECTA"
		/db_xref="taxon:3702"
		/clone="LERGX20"
		/clone_1lb="LERG"
		/note="Organ: Leaf; Vector: pUC19K; Total genomic DNA was sheared to 0.4-0.7 Kbp before ligation."
Query Match		12.2%; Score 84.8; DB 28; Length 613;

Best Local Similarity 97.7%; Pred. No. 0.00012;
Matches 86; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 608 GTTGGCTTTTAACTTCACTGTCAGTGTGATTTTGGTAGTACCTGTG 667
Db 613 GTTGGAGTTTAACTTCACTGTCAGTGTGATTTTGGTAGTACCTGTG 554

QY 668 GTTTCCTTGGCTATAGCTTGAANA 695
Db 553 GTTTCCTTGGCTATAGCTTGAANA 526

RESULT 11
AU236368
LOCUS
DEFINITION AU236368 647 bp mRNA linear EST 01-APR-2002
AU236368 RAFL14 Arabidopsis thaliana cDNA clone RAFL14-93-K05 5',
mRNA sequence.
ACCESSION AU236368
VERSION AU236368.1 GI:19875537
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana

REFERENCE
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 647)
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
Oono, Y., Sekurai, T., Carinci, P., Kawai, J., Itoh, M., Ishii, Y.,
Arakawa, T., Shibata, K., Shingawa, A., Muramatsu, M., Hayashizaki, Y.,
and Shinzaki, K.
Large scale analysis of Arabidopsis full-length cDNA (2002b)
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekirc@riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified Lambda FIC-1 vector (Carinci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/Plant/index_e.html) for further
details.

FEATURES
source
1.647
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/clone="RAFL14-93-K05"
/issue_type="root"
/lab_host="DH10B"
/clone_1ib="RAFL14"
/note="Site_1: BamHI; Site_2: SalI"

ORIGIN
Query Match 12.1%; Score 84; DB 9; Length 647;
Best Local Similarity 100.0%; Pred. No. 0.00015;
Matches 84; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 199 AGTTATTTTCTCCCGAGTCTGACCTCACTACCTCCTCCGCGCTTAACT 258
Db 1 AGTTATTTTCTCCCGAGTCTGACCTCACTACCTCCTCCGCGCTTAACT 60

QY 259 TAGCTTCCTCGTGTACTCTGT 282
Db 61 TAGCTTCCTCGTGTACTCTGT 84

RESULT 12
BZ558518

LOCUS BZ558518 1872 bp DNA linear GSS 17-DEC-2002
DEFINITION pa98401.209.x1 pacs2-164 Pseudomonas aeruginosa genomic clone
pa98401.209, genomic survey sequence.
ACCESSION BZ558518
VERSION BZ558518.1 GI:27173068
KEYWORDS GSS.
SOURCE Pseudomonas aeruginosa
ORGANISM Pseudomonas aeruginosa
Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
Pseudomonadaceae; Pseudomonas.
1 (bases 1 to 1872)
REFERENCE
AUTHORS Spencer, D.H., Raymond, C.K., Smith, E.E., Sims, E.E., Hastings, M.,
Burns, J.L., Kaul, R. and Olsen, M.V.
TITLE Whole-Genome-Sequence variation among multiple isolates of
Pseudomonas aeruginosa library
J. Bacteriol. (2002) In press
COMMENT Contact: Chris K. Raymond
Genome Center
University of Washington
Box 352145, Seattle, WA 98105-2145, USA
Tel: 2062216954
Fax: 2066857244
Email: ckraymond@u.washington.edu
Class: shotgun.

FEATURES
source
1.1872
/organism="Pseudomonas aeruginosa"
/mol_type="genomic DNA"
/strain="2-164"
/db_xref="taxon:287"
/clone="pa98401.209"
/clone_1ib="pacs2-164"
/note="clinical isolate 2-164 Whole genomic shotgun
library."

ORIGIN
Query Match 11.9%; Score 82.8; DB 28; Length 1872;
Best Local Similarity 46.2%; Pred. No. 0.00012;
Matches 304; Conservative 0; Mismatches 348; Indels 6; Gaps 1;

QY 2 ACACACAAACAAATTAAGCAGTCATGTCGATCAATTAATATGCGAACAAT 61
Db 527 AA 586

QY 62 TAAGTTAACGAAAAAGAAAAAGATCAAAAAATCAAAAAACAACTGAATG 121
Db 587 AA 646

QY 122 AAAATTTGAGTCAGATCGGAAAAAGAGCCGTTTAGAGCTTAATAGCTTCCTCA 181
Db 647 AAAAAAGGGGGGAAAAAGAAAAAGGGGGGGGTTTGGGGGATGTTTTTTTTT 706

QY 182 TTTGCTCTTCTGTCAGTTATTTCTCTCTCGGAGCTGACTCACTCTCACT 241
Db 707 TTTTCTTTTGTGTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 766

QY 242 CTCGCGCGCTTAACTTACGTTCCGCTGTTTACTCTGTAAGTTTTCGCTTAGAC 301
Db 767 TGTGTGTTT 826

QY 302 CTCGAGTCGCTCACGCGATCTGTCGATTTCTTTTCTTCGCTGAAAAA 361
Db 827 TTTTCTTTTGTGTTTTGTTTTTTTTTTTGTGTTTTTTTTTTGTTTTGTTGT 886

QY 362 TTGCGCTAAATGTCGATTTGAAGTTTTTGTGCTAGAGGTTATTTTCCCTAAT 421
Db 887 TTGTTTTTTGTTTTTTTTTTTTTTTTTTTTTTGTTTTGTTTTTTGTTTTGTTT 946

QY 422 TTATAGTCTTAGTAACGATACCGGCTTACGTTTTTGTTCATTTTGTGCTT 481
Db 947 TTTTATTTGTTTTGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1006

QY 482 TCACGTTTAGTCGATCGAGATTTGACTGTGAAAAATCCTTCGTTTTTGGTTT 541

D_b 1007 TTTTCTTTTTATTTTGTGTAATTTTGTTTAAAGATTTTTTTTTAA 1066
O_y 542 TGTTTCATAAATC-----GGATGATCAACCTTTGCGCTTGANGTTGTTTTTG 595
D_b 1067 TAATTGTATATTTTTTTTTGTTTTTAAATTTTGGTTTTTAAATTTTTTTTTT 1126
O_y 596 AGCCATGCGTGTGGCTGTGTAAACTGCACGTCATGATGCTGATTTNAGATTTT 653
D_b 1127 TTTTCTTGTGTTGTTTTTTTTTTTTTTTTTTTTTTTTTGGTTTGGTT 1184

RESULT 13	CNS00ETV	1068 bp	DNA	linear	GSS 04-JUN-1999
LOCUS	CNS00ETV				
DEFINITION	Drosophila melanogaster genome survey sequence 17 end of BAC: BACR29P07 of RP11-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.				
ACCESSION	AL069846				
VERSION	AL069846.1	GI:4949798			
KEYWORDS	GSS.				
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 1068)				
AUTHORS	Genoscope.				
TITLE	Direct Submission				
JOURNAL	Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :				

```

FEATURES
source
location/Qualifiers
1..1068
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone_1b="BACR29P07"
/clone_1b="RPCI-98"
/note="end : 17"
ORIGIN

```

	Query Match	Similarity	11.7%	Score 81.2	DB 29	Length 1068
	Best Local	Similarity	35.6%	Pred. No. 0.00031		
	Matches	245	Conservative	82	Mismatches 361	Indels 0; Gaps 0
QY	2	ACAAACAAACAAATTTAGCAATGCTGTCAGCAATTAATTAATAGTGGAAACAT	61			
Db	248	AAA	307			
QY	62	TAAAGTTAGCGAAAAAGAAAAAGATCAAAAAATGAAAAAAATCAATCTGAATG	121			
Db	308	AAA	367			
QY	122	AAATTTGAGTCCAGATCGAATAACGAGCCGCTTTGAGCTTAATAGTTCCTCA	181			
Db	368	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAACCCTTTTCCTATATATATATTTTAA	427			

[illegible][illegible]

FEATURES
SOURCE

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 19:34:49 ; Search time 35.059 Seconds
(without alignments)
11001.196 Million cell updates/sec

Title: US-10-030-829-1_COPY_1_695
Perfect score: 695
Sequence: 1 gccaacaacaacaataa.....ggctgctatagtgtraaa 695

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA:*
1: /cgn2_6/prodata/2/ina/5A COMB.seq:*
2: /cgn2_6/prodata/2/ina/5B COMB.seq:*
3: /cgn2_6/prodata/2/ina/6A COMB.seq:*
4: /cgn2_6/prodata/2/ina/6B COMB.seq:*
5: /cgn2_6/prodata/2/ina/6CTUS COMB.seq:*
6: /cgn2_6/prodata/2/ina/backfillseq.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	85.6	12.3	7218	1	US-08-232-463-14 Sequence 14, Appl
2	52.6	7.6	7286	3	US-09-331-581-3 Sequence 3, Appl1
3	52.6	7.6	7938	4	US-09-331-581-14 Sequence 14, Appl
4	50.6	7.3	396	4	US-09-640-173-53 Sequence 53, Appl
5	50.6	7.3	396	4	US-09-713-550-53 Sequence 53, Appl
6	50.6	7.3	10619	4	US-10-204-708-4 Sequence 4, Appl1
7	50.2	7.2	5562	4	US-10-204-708-63 Sequence 63, Appl
8	49.4	7.1	19124	2	US-08-487-826B-13 Sequence 13, Appl
9	47.8	6.9	2447	2	US-09-014-969-14 Sequence 14, Appl
10	47.4	6.8	5666	4	US-10-204-708-29 Sequence 29, Appl
11	47.4	6.8	8607	4	US-10-204-708-72 Sequence 72, Appl
12	46.8	6.7	19233	4	US-10-204-708-45 Sequence 45, Appl
13	46.8	6.7	19233	4	US-10-204-708-45 Sequence 45, Appl
14	45.6	6.6	7218	1	US-08-232-463-14 Sequence 14, Appl
15	45.6	6.6	289	3	US-09-007-005-17 Sequence 17, Appl
16	45.6	6.6	289	3	US-09-244-796-17 Sequence 17, Appl
17	45.4	6.5	11049	4	US-10-204-708-23 Sequence 23, Appl
18	44.2	6.4	1493	1	US-08-340-820-24 Sequence 24, Appl
19	44.2	6.4	1493	1	US-08-593-535-24 Sequence 24, Appl
20	44.2	6.4	240	1	US-08-628-417-6 Sequence 6, Appl1
21	43.6	6.3	2394	4	US-09-800-729-33 Sequence 33, Appl
22	43.6	6.3	359	4	US-09-621-976-16008 Sequence 16008, A
23	43.6	6.3	2915	4	US-09-336-115C-5 Sequence 5, Appl1
24	43.6	6.3	5501	4	US-10-204-708-38 Sequence 38, Appl
25	43.4	6.2	140	1	US-08-628-417-5 Sequence 5, Appl1
26	43.4	6.2	359	4	US-09-621-976-16019 Sequence 16019, A
27	43.4	6.2	362	4	US-09-621-976-16010 Sequence 16010, A

28	43.4	6.2	365	4	US-09-621-976-16042 Sequence 16042, A
29	43.4	6.2	5455	4	US-10-204-708-33 Sequence 33, Appl
30	43.4	6.2	11015	4	US-10-204-708-55 Sequence 55, Appl
31	43	6.2	578	4	US-09-602-877A-95 Sequence 95, Appl
32	43	6.2	8607	4	US-10-204-708-72 Sequence 72, Appl
33	43	6.2	8961	4	US-10-204-708-80 Sequence 80, Appl
34	43	6.2	19124	2	US-08-487-826B-13 Sequence 13, Appl
35	42.8	6.2	1114	4	US-09-152-060-41 Sequence 41, Appl
36	42.8	6.2	11050	4	US-10-204-708-85 Sequence 85, Appl
37	42.8	6.2	11131	4	US-10-204-708-27 Sequence 27, Appl
38	42.6	6.1	240	1	US-08-628-417-6 Sequence 6, Appl1
39	42.6	6.1	371	4	US-09-621-976-16048 Sequence 16048, A
40	42.6	6.1	10144	4	US-10-204-708-94 Sequence 94, Appl
41	42.4	6.1	270	2	US-08-520-678A-30 Sequence 30, Appl
42	42.4	6.1	270	3	US-08-897-126-30 Sequence 89, Appl
43	42.2	6.1	5844	4	US-10-204-708-89 Sequence 89, Appl
44	42.2	6.1	6866	4	US-10-204-708-20 Sequence 20, Appl
45	42	6.0	357	4	US-09-621-976-16058 Sequence 16058, A

ALIGNMENTS

RESULT 1
US-08-232-463-14
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
TITLE OF INVENTION: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
CLASSIFICATION: 435
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PT29PT-F18
US-08-232-463-14
Query Match 12.3%; Score 85.6; DB 1; Length 7218;

Best Local Similarity 6.4%; Pred. No. 1,4e-12;
Matches 25; Conservative 232; Mismatches 131; Indels 0; Gaps 0

```

: RESULT 2
: US-09-331-581-3
: Sequence 3, Application US/09331581
: Patent No. 6130070
: GENERAL INFORMATION:
: APPLICANT: TOMDA, Hideki
: APPLICANT: HAMA, Yuko
: APPLICANT: KUNAGAI, Hiromi:cho
: TITLE OF INVENTION: INDICATION PROMOTER GENE AND SECRETORY SIGNAL GENE
: TITLE OF INVENTION: USALE IN SCHIZOSACCHAROMYCES POMBE, EXPRESSION VECTORS
: TITLE OF INVENTION: HAVING THE SAME, AND USE THEREOF
: FILE REFERENCE: 0059-1142-0PCT
: CURRENT APPLICATION NUMBER: US/09/331,581
: CURRENT FILING DATE: 1999-06-30
: EARLIER APPLICATION NUMBER: PCT/JP98/04929
: EARLIER FILING DATE: 1998-10-30
: EARLIER APPLICATION NUMBER: JP 9-314608
: EARLIER FILING DATE: 1997-10-31
: NUMBER OF SEQ ID NOS: 24
: SOFTWARE: PatentIn Ver. 2.1
: SEQ ID NO 3
: LENGTH: 7286
: TYPE: DNA
: ORGANISM: Artificial Sequence
: FEATURE:
: OTHER INFORMATION: Description of Artificial Sequence:DNA
: US-09-331-581-3

```

Query Match	7.6%	Score 52.6;	DB 3;	Length 7286;
Best Local Similarity	56.7%;	Pred. No. 0.00044;		
Matches 97;	Conservative	0;	Mismatches 74;	Indels 0;
				Gaps 0;

[illegible]

Db 4679 TACCTTCTGAGCGGAAAGAACCAAGCCGATCAGACATGATTAAGATACAT 4723

```

RESULT 3
US-09-331-561-14
; Sequence 14, Application US/09331581
; Patent No. 6130070
; GENERAL INFORMATION:
; APPLICANT: TOMDA, Hideki
; APPLICANT: HAMA, Yuko
; APPLICANT: KUMAGAI, Hiromichi
; TITLE OF INVENTION: INDUCTION PROMOTER GENE AND SECRETORY SIGNAL GENE
; TITLE OF INVENTION: USABLE IN SCHIZOSACCHAROMYCES POMBE, EXPRESSION VECTORS
; TITLE OF INVENTION: HAVING THE SAME, AND USE THEREOF
; FILE REFERENCE: 0059-1142-0PCT
; CURRENT APPLICATION NUMBER: US/09/331,581
; CURRENT FILING DATE: 1999-06-30
; EARLIER APPLICATION NUMBER: PCT/J998/04929
; EARLIER FILING DATE: 1998-10-30
; EARLIER APPLICATION NUMBER: JP 9-314608
; EARLIER FILING DATE: 1997-10-31
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 14
; LENGTH: 7938
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:DNA
US-09-331-561-14

```

Query Match	7.6%;	Score 52.6;	DB 3;	Length 7938;
Best Local Similarity	56.7%;	Pred. No. 0.00045;		
Matches. 97;	Conservative	0;	Mismatches 74;	Indels 0;
				Gaps 0;

QY 9 AACAAAAATTAGCAAGCATGTCGTCAGCAATAAATTAATAGCGGAAACATTTAGTTA 68
Db 5211 AACCAATCATGAGAAAGATGTCTATGATGCTGAAATATTAATGACGTCACAAAGCAAAA 5270
QY 69 AGCGAAAAAGGAAAAAAGGTACAATAATGAAAAAATACTCAACTGAATGAAAAATT 128
Db 5271 AAG 5330
QY 129 GAGATCGAATCGAAAAACGAGCCGTTTTAGAGTTAATAGCTTCT 179
Db 5331 TACCTTCGAGCGCGAAAGAACGAGCCGATTCAGACATGATTAGATCAT 5381

```

RESULT 4
US-09-640-173-53
Sequence 53: Application US/09640173
Patent No. 6613515
GENERAL INFORMATION:
APPLICANT: Xu, Jiangchun
APPLICANT: Stolk, John A.
TITLE OF INVENTION: OVARIAN TUMOR SEQUENCES AND
TITLE OF INVENTION: METHODS OF USE THEREFOR
FILE REFERENCE: 210121.484C2
CURRENT APPLICATION NUMBER: US/09/640,173
CURRENT FILING DATE: 2000-08-15
NUMBER OF SEQ ID NOS: 196
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 53
LENGTH: 396
TYPE: DNA
ORGANISM: Homo sapien
FEATURES:
NAME/KEY: misc feature
LOCATION: (1)..(396)
OTHER INFORMATION: n = A,T,C or G
US-09-640-173-53

```

Query Match 7.3%; Score 50.6; DB 4; Length 396;
Best Local Similarity 49.8%; Pred. No. 0.00049;
Matches 119; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 389 TTTTGTGCTAGGGTACTTTTCCCTATATTAAGTCTTAGAGTAACATACCTG 448
DB 18 TTTTGTGCTAGGGTACTTTTCCCTATATTAAGTCTTAGAGTAACATACCTG 77

QY 449 CGCTTACGCTTTGCTCATTTGTTGCTTCCGTTAGTGGCGATGAGATAT 508
DB 78 TTTTGTGCTAGGGTACTTTTCCCTATATTAAGTCTTAGAGTAACATACCTG 137

QY 509 TTGACTGTGAAAAATCCTGCTTTTGGTTTGTTCATATAAATCGATTGATCTAC 568
DB 138 TTTTGTGCTAGGGTACTTTTCCCTATATTAAGTCTTAGAGTAACATACCTG 197

QY 569 CTTTGTGCTTGAATGTTGTTTGGAGCTTATGCTTGGCTGTGATTAACCTCA 627
DB 198 TTTTGTGCTTGAATGTTGTTTGGAGCTTATGCTTGGCTGTGATTAACCTCA 256

RESULT 5

US-09-713-550-53
; Sequence 53, Application US/09713550
; Patent No. 6617109
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; FILE REFERENCE: 210121.484C4
; CURRENT APPLICATION NUMBER: US/09/713,550
; NUMBER OF SEQ ID NOS: 205
; SOFTWARE: FaelsSO for Windows Version 3.0
; SEQ ID NO 53
; LENGTH: 396
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(396)
; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-53

Query Match 7.3%; Score 50.6; DB 4; Length 396;
Best Local Similarity 49.8%; Pred. No. 0.00049;
Matches 119; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

QY 389 TTTTGTGCTAGGGTACTTTTCCCTATATTAAGTCTTAGAGTAACATACCTG 448
DB 18 TTTTGTGCTAGGGTACTTTTCCCTATATTAAGTCTTAGAGTAACATACCTG 77

QY 449 CGCTTACGCTTTGCTCATTTGTTGCTTCCGTTAGTGGCGATGAGATAT 508
DB 78 TTTTGTGCTAGGGTACTTTTCCCTATATTAAGTCTTAGAGTAACATACCTG 137

QY 509 TTGACTGTGAAAAATCCTGCTTTTGGTTTGTTCATATAAATCGATTGATCTAC 568
DB 138 TTTTGTGCTAGGGTACTTTTCCCTATATTAAGTCTTAGAGTAACATACCTG 197

QY 569 CTTTGTGCTTGAATGTTGTTTGGAGCTTATGCTTGGCTGTGATTAACCTCA 627
DB 198 TTTTGTGCTTGAATGTTGTTTGGAGCTTATGCTTGGCTGTGATTAACCTCA 256

RESULT 6

US-10-204-708-4
; Sequence 4, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian

APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 4
; LENGTH: 10619
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-4

Query Match 7.3%; Score 50.6; DB 4; Length 10619;
Best Local Similarity 49.8%; Pred. No. 0.0017;
Matches 128; Conservative 0; Mismatches 129; Indels 0; Gaps 0;

QY 404 TTACTTTTCCCTATATTAATGTTCTTAGTAACATACCTGCTTACTGTTTTT 463
DB 9137 TTTTATTTATTTGTTATTAATGATTAATTTTATTTTATTTTATTTTATTTT 9196

QY 464 GTTCAATTTGTTGCTTTCACCGTTTATGCTGATCGAGATTGACTGAAAAAT 523
DB 9197 TTTGCTTTGTTTTTTTTTTTTTTTTTTTTTTCGTTTGAATATTTTGTATATT 9256

QY 524 CCTGCTTTTGGCTTTTGTTCATATAAATCGATTGATCCTTTGCTTGTAT 583
DB 9257 AGTTTTTTTGGTTGGAATGATGATTTTGTTTTGGATTTTATTTTATGTAAGA 9316

QY 584 GTTGTGTTTTTGGCTTATGCTTGTGCTTGTATTAATCACTGATGATGAT 643
DB 9317 GTTTTGTGTTTGGTGTGTTTATTTATTTGATGATGATTTATTTATTTTATTT 9376

QY 644 TTGAGATTTTGTAGTG 660
DB 9377 TGTATTTATTTGTTAGTG 9393

RESULT 7

US-10-204-708-63
; Sequence 63, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; PRIOR FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01

NUMBER OF SEQ ID NOS: 98
SEQ ID NO: 63
LENGTH: 5562
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically created genomic DNA (Homo sapiens)
US-08-204-708-63

Query Match 7.2%; Score 50.2; DB 4; Length 5562;
Best Local Similarity 48.1%; Pred. No. 0.0017;
Matches 142; Conservative 0; Mismatches 153; Indels 0; Gaps 0;

QY 390 TTTTGCTAGGAGTACTTTTCCCTATTTTATTTAGTACGATCTCG 449
DB 1394 TATTTATTTATTTATTTTTCGTTAGAAATGATTTTGTATTTAT 1453
QY 450 GTCTTACTGTTTGTCTATTTTGTGCTTACCGTTAGCTGAGTAT 509
DB 1454 ATTTTATTTATTTAGTTAGTTTATTTTATTTTATTTATTTATTT 1513
QY 510 TGAAGTGAATAATCTGCTTTTGTGTTTGTATTAATGATGATCTAC 569
DB 1514 TGTATTTAGATTTTATTTTATTTTATTTTATTTTATTTTATTT 1573
QY 570 TTTTGCTTGAATTTTGTGAGCTTATGCGTTGCTGTTAATCTCAG 629
DB 1574 TTTTGTGTTTGAAGATTTATTTATTTTATTTTATTTTATTTT 1633
QY 630 TTCATGTCGATTTTGAATTTTGTGATGCTGCTGCTTCTTGTGCT 684
DB 1634 GATAGGTATTTTGAATTTTGTGATTAATGATTTTATTTATTTAT 1688

RESULT 8

US-08-487-826B-13/C
Sequence 13, Application US/08487826B
Patent No. 5993827

GENERAL INFORMATION:

APPLICANT: Sim, Kim L.
APPLICANT: Chitnis, Chetan
APPLICANT: Miller, Louis H.
APPLICANT: Peterson, David S.
APPLICANT: Su, Xin-zhaun
APPLICANT: Wellens, Thomas E.
TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
NUMBER OF SEQUENCES: 45
CORRESPONDENCE ADDRESS:
ADDRESSEE: Knobbe Martens Olson & Bear
STREET: 620 Newport Center Drive 16th Floor
CITY: Newport Beach
STATE: California
COUNTRY: US
ZIP: 92660

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/487,826B
FILING DATE: 10-SEP-1993
CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: Israelien, Ned
REGISTRATION NUMBER: 29,655
REFERENCE/DOCKET NUMBER: NIH121,001CP1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (619) 235-8550
TELEFAX: (619) 235-0176
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:

LENGTH: 19124 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULAR TYPE: CDNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
US-08-487-826B-13

Query Match 7.1%; Score 49.4; DB 2; Length 19124;
Best Local Similarity 47.9%; Pred. No. 0.0042;
Matches 172; Conservative 0; Mismatches 186; Indels 1; Gaps 1;

QY 318 GCATGATCTGTCGATCTGATTTCTTTTCTTCGCGAAATATGCCATGTTCTC 377
DB 15976 GCATGATTAATTTATTTATTTTATTTTATTTATTTATTTATTTTAT 15917
QY 378 GATTCGAAGTTTGTGCTAATGCGTAACTTTTCCCTATATTTATGTTCTAG 437
DB 15916 TCATTAATTTTATTTTATTTTATTTTATTTATTTATTTATTTATTT 15857
QY 438 AACGATCTGCTGCTTACTGTTTGTTCATTTTGTGTGCTTCACCGTTAGCGCT 497
DB 15856 TTAATTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTT 15797
QY 498 GATCGAGATTTGACCTGGAATAATCTGTTTGTGTTTGTTCATTAATCG 557
DB 15796 TATGATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTATTT 15737
QY 558 G-ATTGATCTACCTTTTGTGCTTGAATTTGTTTGTGAGCTAATGCTTGTGCT 616
DB 15736 TATTTTATTAATTTTATTTTATTTTATTTTATTTTATTTTATTT 15677
QY 617 TATTAATCTACGCTCAGCTGATTTGATTTGATTTGTTAGTACGCTGCTTCT 675
DB 15676 AATGTTTATTTTCTTCTTGTGTTTATTTTATTTTATTTTATTTATTT 15618

RESULT 9

US-09-014-969-14/C
Sequence 14, Application US/09014969
Patent No. 5965397

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
APPLICANT: McCoy, John M.
APPLICANT: Lavallee, Edward R.
APPLICANT: Racie, Lisa A.
APPLICANT: Werberg, David
APPLICANT: Treacy, Maurice
APPLICANT: Spaulding, Vikki
APPLICANT: Agostino, Michael J.
TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
TITLE OF INVENTION: ENCODING THEM
NUMBER OF SEQUENCES: 32
CORRESPONDENCE ADDRESS:
ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/014,969
FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323

TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2447 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-014-969-14

Query Match 6.9%; Score 47.8; DB 2; Length 2447;
Best Local Similarity 47.2%; Pred. No. 0.0051;
Matches 142; Conservative 1; Mismatches 158; Indels 0; Gaps 0;

QY 390 TTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGGTAACGATACCTGC 449
DB 2447 TTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGGTAACGATACCTGC 2388
QY 450 GCTTACGTTTGTGTTCAATTTGTTGCTTACCGTTAGTCGTAACGAGATTT 509
DB 2387 TTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGGTAACGATACCTGC 2328
QY 510 TGACTGTGAAAAATCCCTGTTTGTGTTTGTTCATATTAATCGATTGATCTACC 569
DB 2327 TTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGGTAACGATACCTGC 2268
QY 570 TTTTGTGCTTGTATGTTTGTGTTTGTGACCTATGCGTTGCTGTATTAACCTCAGC 629
DB 2267 TTTTGTGCTTGTATGTTTGTGTTTGTGACCTATGCGTTGCTGTATTAACCTCAGC 2208
QY 630 TTCATGTGATGATTTGAGATTTTGGTAGTACGTTGGTTTGTGCGCTATAGTT 689
DB 2207 ATCAATACACATCCAGTATGTAACCATATATACATATATACAGCCATTATTTA 2148
QY 690 G 690
DB 2147 G 2147

RESULT 10
US-10-204-708-29
Sequence 29, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBERCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204, 708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98

SEQ ID NO 29
LENGTH: 5666
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-29

Query Match 6.8%; Score 47.4; DB 4; Length 5666;
Best Local Similarity 49.2%; Pred. No. 0.0088;
Matches 153; Conservative 0; Mismatches 156; Indels 2; Gaps 1;

QY 384 GAAGTTTGTGCTATGGGTTACTTTTCCCTATTTTATGTTCTTAGGTAACGAT 443
DB 3358 GATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3417
QY 444 ACCGCGCTTACGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 503
DB 3418 TTTTACGTTATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 3477
QY 504 AGTATTTGACTGTGAAAAATCCCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 561
DB 3478 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 3537
QY 562 GATCTACCTTTGTGCTTGTGATGTTTGTGTTTGTGACCTATGCGTTGCTGTATTA 621
DB 3538 GGTGTTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 3597
QY 622 ACTTACGTTTCAATGCTGATTTTGTGATTTTGTGATGATCTGCTGCTTCTTTGGTGC 681
DB 3598 TGTGTTTGTGAAAGATATGATTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 3657
QY 682 TATAGTTGTA 692
DB 3658 TATGTTTATTA 3668

RESULT 11
US-10-204-708-72
Sequence 72, Application US/10204708
Patent No. 6677731
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBERCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
FILE REFERENCE: 5013.1012
CURRENT APPLICATION NUMBER: US/10/204, 708
CURRENT FILING DATE: 2003-05-06
PRIOR APPLICATION NUMBER: PCT/EP01/03971
PRIOR FILING DATE: 2001-04-06
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06
PRIOR APPLICATION NUMBER: DE 10019173.8
PRIOR FILING DATE: 2000-04-07
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 98
SEQ ID NO 72
LENGTH: 8607
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-72

Query Match 6.8%; Score 47.4; DB 4; Length 8607;
Best Local Similarity 46.7%; Pred. No. 0.01;
Matches 150; Conservative 0; Mismatches 171; Indels 0; Gaps 0;

QY 374 TCTGATTTGGAAGGTTTGTGCTATGGGTTACTTTTCCCTATTTTATGATTTCTT 433
DB 6966 TATGTTTATTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 7025
QY 434 AGCTAACGATACCTGCGTTTACGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTG 493
DB 7026 TTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 7085

QY 494 CGCTGATCGAGTATTGACGTGAAAAATCCTTCTTTTGGTTTGGTTCATATA 553
 Db 7086 TTTTGTGTTTATTTTGTGTTTATATATTTTATAGATGTTGTTTGGAGAA 7145
 QY 554 ATCGGATTATCTACCTTTTGGCTTGATGTTTGGTTCATAGCGTTGGC 613
 Db 7146 ATAAAAGAGTTTATGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 7205
 QY 614 TTGTATTAATTCACCTTCATGCTGATTTGAGATTGTTGGTATGCTGGTTCT 673
 Db 7206 AGAATTAGTATTTTGTGATGATGTTGTTGACGTGAAAAAGGATATGATTTTA 7265
 QY 674 TTGGTGGCTATAGTTGTAAA 694
 Db 7266 AGTATGTTTATTTTGTGTTAA 7286

RESULT 12
 US-10-204-708-79
 ; Sequence 79, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 79
 ; LENGTH: 8961
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; NAME/KEY: unsure
 ; LOCATION: (3866)
 ; OTHER INFORMATION: n is a or g or c or t
 US-10-204-708-79

Query Match 6.8%; Score 47; DB 4; Length 8961;
 Best Local Similarity 48.9%; Pred. No. 0.013;
 Matches 155; Conservative 0; Mismatches 160; Indels 2; Gaps 1;
 QY 341 TCTTTTCTTCGCTGAAAAATGCTTAATGTTTCGATTTGGAAGTTTGGCTAT 400
 Db 6013 TATTTGTGATGAGGGAATGTTGTTTATTTTAAAGATGATGTTAGTTT 6072
 QY 401 GGGTACTTTTCCCTAATTTTATAGTCTTAGGTAACGATACCTGCTTACTGTT 460
 Db 6073 TAAGTAAGTATTTTATTTTGTGTTTATTTAGTTTATTTAGTTATTTATTT 6132
 QY 461 TTGTTCATTTTGTGCTTTCACCGTTAGTCCGTCGATCGAGATTTGACTGGAA 520
 Db 6133 ATTTTGTGATTTTATGTTTATTTAGTTTATTTTATTTTATTTATTTTATTT 6190
 QY 521 AATCTTGTGTTTGTGTTTGTTCATATAAAGCATGATCTACCTTTGGCTTT 580
 Db 6191 TTTTGTGTTTATTTTATTTTATTTTATTTGAGTTTGTGTTTGTGTTTGTGTTT 6250

QY 581 GATGTTGTTTGGAGCTATGCGTTGTTGGCTTTATATACCTTCAGTTGATG 640
 Db 6251 ATTTGATATTTATTTTATTTTGTGTTTATTTTATTTTATTTAGTGAAGTTGAGTGT 6310
 QY 641 ATTTGAGATTTTGGTA 657
 Db 6311 TGGGAGAGATGTTTATTTTA 6327

RESULT 13
 US-10-204-708-45
 ; Sequence 45, Application US/10204708
 ; Patent No. 6677731
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PIEPENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
 ; FILE REFERENCE: 5013.1012
 ; CURRENT APPLICATION NUMBER: US/10/204,708
 ; PRIOR FILING DATE: 2003-05-06
 ; PRIOR APPLICATION NUMBER: PCT/EP01/03971
 ; PRIOR FILING DATE: 2001-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019058.8
 ; PRIOR FILING DATE: 2000-04-06
 ; PRIOR APPLICATION NUMBER: DE 10019173.8
 ; PRIOR FILING DATE: 2000-04-07
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 98
 ; SEQ ID NO 45
 ; LENGTH: 19233
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 ; NAME/KEY: unsure
 ; LOCATION: (76, 178..179, 273, 586, 648, 651, 920, 1014, 1173, 1197, 1228)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1264, 1267, 1559, 1607, 1677, 1703, 1742, 1848, 1850, 2033)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (2168, 2254, 2257, 2280, 2291, 2723, 3018, 3080, 4205, 4209)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (4216, 4225, 4239, 4246, 4254, 4265, 4270, 4274, 4288, 4294..4295)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (4298, 4363, 4610, 4627, 4675, 4693, 4698, 4738, 4777, 4781)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (4787, 4825, 4829, 4838, 4855, 5462, 5494, 5496, 5500, 5674)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (5690, 5697, 5705, 5715, 5890, 5910, 5913, 6027, 6080, 6109)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (6244, 6288, 6776, 6873, 6891, 6980, 7028, 7656, 7876, 8030)
 ; OTHER INFORMATION: n is a or g or c or t
 ; FEATURE:
 ; NAME/KEY: unsure

LOCATION: (8059, 8075, 8632, 8636, 9700, 9717, 9815, 9818, 9867, 9879)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (10005, 10075, 10150, 10159, 10165, 10356, 11123, 11166, 11360)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (11703, 11933, 12050, 12340, 12981, 13564, 13764, 13765, 13787)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (13846, 14168, 16566, 17006, 17217, 17650, 17653, 17656, 17898)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (18130, 18143, 18151, 18188, 18498, 18549, 18936, 18946, 19165)
OTHER INFORMATION: n is a or g or c or t
FEATURE:
NAME/KEY: unsure
LOCATION: (19200)
OTHER INFORMATION: n is a or g or c or t
US-10-204-708-45

Query Match 6.7%; Score 46.8; DB 4; Length 19233;
Best Local Similarity 48.5%; Pred. No. 0.02;
Matches 129; Conservative 0; Mismatches 137; Indels 0; Gaps 0;

QY 360 AATGCCCTAATGCTTCGATTTGCAAGGTTTGGCTAGTGGCTACTTTTCCCTAT 419
DB 11391 AATTATATGTTTGTGGAGATATAGAAATTTTATTAACGTAAATTTTGTGTT 11450
QY 420 ATTATAGTCTTAGGTAACGATACCTGCTTACTGTTTGTTCATTGTTGTGTC 479
DB 11451 TTTTATATTTTATTTTATTTTATTTAGTAATGCTGTTTGTGTTTATTTGTAAGTT 11510
QY 480 TTTCACGCTTATGCTGCTGATCGAGATTTGACTGTAATAATCCCTCGTTTGTGTT 539
DB 11511 AGTTTGTGTTTATTAAGATTTTATATATGAATTAATAATATTTATTTTGTGTG 11570
QY 540 TTGTTTCAATATAATCGATGATGATCTACCTTTGTGCTTGTGATGTTTGTGTTGAGCC 599
DB 11571 TTGGTTTTTTTATGATGTTTTTTTTTTTTTTTATAGAGTGGGTTTTGTTATTTAGTT 11630
QY 600 TATGCTGTGCTGCTGTATACCT 625
DB 11631 GGAGTGTAGCATATATGTTATGTTT 11656

RESULT 14

US-08-232-463-14/C
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHREIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463

FILING DATE:
CLASSIFICATION: 435
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
FILING DATE: 26-AUG-1991
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMM
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 836-9300
TELEFAX: (703) 683-4109
TELEX: 899149
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: pTZ9pt-Fls
US-08-232-463-14

Query Match 6.6%; Score 46; DB 1; Length 7218;
Best Local Similarity 5.6%; Pred. No. 0.022;
Matches 10; Conservative 115; Mismatches 55; Indels 0; Gaps 0;

QY 1 GACAAACAACAATAATGACAGTCTGTCGATTAATAATTAATGAGGACAACA 60
DB 1231 RRR 1172
QY 61 TTAAGTAAAGCAAGAAAGAAAGAAAGTACAAATAAGAAACAAATCAACTGAAT 120
DB 1171 RRR 1112
QY 121 GAAATTTGAGCGCAATCGAAGAAACAGAGCGTTTATAGCTTAATAGCTCTC 180
DB 1111 RRR 1052

RESULT 15

US-09-007-005-17/C
Sequence 17, Application US/09007005B
Patent No. 6258358
GENERAL INFORMATION:
APPLICANT: Szostak, Jack W.
APPLICANT: Roberts, Richard W.
APPLICANT: Liu, Rih
TITLE OF INVENTION: SELECTION OF PROTEINS USING RNA-PROTEIN
TITLE OF INVENTION: FUSIONS
FILE REFERENCE: 00786/350003
CURRENT APPLICATION NUMBER: US/09/007,005B
CURRENT FILING DATE: 1998-01-14
EARLIER APPLICATION NUMBER: 60/035,963
EARLIER FILING DATE: 1997-01-27
EARLIER APPLICATION NUMBER: 60/064,491
NUMBER OF SEQ ID NOS: 33
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 17
LENGTH: 289
TYPE: RNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Translation template
NAME/KEY: misc_feature
LOCATION: (1)...(289)
OTHER INFORMATION: n = A,T,C or G
US-09-007-005-17

QY	260	ACGCTTCGCGTGGTACCTCTGTAAGTTTTCGCTTAAAGCCTCGATGCGCTACCGC	319
Db	971629	TT	971688
QY	320	ATGCATCTGCTCGATTTCTCTTTTCTGCGAAAAATGGCCATATGCTCTCGA	379
Db	971689	TT	971748
QY	380	TTTCGAAGTTTTTGCGCTATGGGTACTTTTCCCTATATTATAGTCTTAGATA	439
Db	971749	TT	971808
QY	440	CGATACCTGCGTCTACTGTTTTTGTCATTGTTGTGCTTCACCGTTAGCGCTGA	499
Db	971809	TTTATTTTTTTTTTTT	971868
QY	500	TCGAGATATTGACTGTGAAAAATCCTCGTTTTTGTTGTTCAATATAATCGA	559
Db	971869	TTTTTTTTTTTTTTTTTTTAAATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	971928
QY	560	TTGATCTACCTTTTGCTCTTGATGTTGTTTTTGAGCCATATGCGTTGGCTGTTA	619
Db	971929	TT	971988
QY	620	TAACTTCAGTTCAGTGTGGAATTGAGATTTTGTTAGTACTGCGTTCCTTGG	677
Db	971989	TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTGGAGATAGGATTTGTTGGTTAG	972046

```

RESULT 2
US-10-311-455-1670
: Sequence 1670, Application US/10311455
: Publication NO. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining the Cytosine Methylation Level of a Cytosine
TITLE OF INVENTION: Cytosine methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1670
LENGTH: 6668
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:
NAME/KEY: unsure
LOCATION: 1936
OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1670

```

	Query Match	Best Local Similarity	Score	DB	Length	Mismatches	Conservative	0;	Indels	Gaps
Qy	156	GTTTTAAGCTTAATAAGCTTCCTCATTTGCTCTCTTCGCAATTATTTCTCTC	215							
Db	2742	GTTTTGTGTTGTT	2801							
Qy	216	CGAGTCCTGACTACTACTCTCACTCCGAGCGCTTAACCTAACGTTCCGCGTTT	275							
Db	2802	TTTTTTTTTTTTTTGTTT	2861							

[illegible]

```

RESULT 3
US-10-239-676-52
/ Sequence 52, Application US/10239676
/ Publication No. US20030082609A1
/ GENERAL INFORMATION:
/ APPLICANT: OLEK, Alexander
/ APPLICANT: PIEPENBROCK, Christian
/ APPLICANT: BERLIN, Kurt
/ TITLE OF INVENTION: Diagnosis of Diseases Associated with Gene Regulation
/ FILE REFERENCE: 5013.1003
/ CURRENT APPLICATION NUMBER: US/10/239,676
/ CURRENT FILING DATE: 2002-09-24
/ PRIOR APPLICATION NUMBER: PCT/EP01/03968
/ DE 10019058.8
/ DE 10019173.8
/ DE 10032529.7
/ DE 10043826.1
/ PRIOR FILING DATE: 2001-04-06
/ 2000-04-06
/ 2000-04-07
/ 2000-06-30
/ 2000-09-01
/ NUMBER OF SEQ ID NOS: 228
/ SEQ ID NO 52
/ LENGTH: 9539
/ TYPE: DNA
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
/ US-10-239-676-52

```

Query Match	9.64;	Score 66.6;	DB 14;	Length 9539;
Best Local Similarity	46.94;	Pred. No. 3.1e-05;		
Matches 207;	Conservative	0;	Mismatches 234;	Indels 0; Gaps 0;
QY	251	TTTAACTTACGTTCTCCGCGTTTACTCGTAAGTTTTCGCTTACAGAGCTCCGATCG	310	
Db	202	TTTTTTTTTTTGATTTTTTTCGTTTATTTTTTGAGCTTTTTTTTTTATTTTTTTTTTC	261	
QY	311	CCTACCGCAGCATTCGTGCTCGAATTCCTTTTCTTCGCTGAAAAATTGCCCTAA	370	
Db	262	GTTTTTTTTTTTTTGCGTTTGTTTTTTTTTTTTTTTTTTTTTCGTTTTTTTTTATTT	321	
QY	371	TGTTTCGATTCGAAAGCTTTTTTGTCGATAGGCTTATTTTCCCAATATTTATAGTT	430	

235

RESULT 8

RESULT 8
US-10-311-455-394
; Sequence 394, Application US/10311455

Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
PRIOR FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 394
LENGTH: 6079
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-394

Query Match
Best Local Similarity 9.4%; Score 65.4; DB 14; Length 6079;
Matches 156; Conservative 0; Mismatches 151; Indels 0; Gaps 0;

Db
371 TGTCTCGATTGCGAGTTTGTGCTGATGCTTCTTTTCCCTATATTATAGTT 430
4235 TTTTCTTTTGAAGATGTTAAGATGTTAGTTGATTTATATATATTTTATTT 4294
Qy 431 CTTAGTAAAGATGCTGCTGCTTCTTCTTTGTTGATTTGTTGCTTCACTGTT 490
4295 TTTTGGAGTTTATTTATTTTGTGAATTTTGTGTTGTTGTTGTTGTTGTTGTT 4354
Qy 491 AGTCGCTGATCGAGTATTGACTGTGAAATCCTTGGTTTGGTTTGGTTTGCATTA 550
4355 GATTAGAAATATATGTTTAACTATTAATTTTGTGTTGTTGTTGTTGTTGTTG 4414
Qy 551 TAAATCGATGATGATCTTCTTGTGCTTGTGATGTTGTTTGTGAGCCATGCTGTT 610
4415 TTAAGTATTTATGTTATTTGAATTTTAAAGATTTTAAATTTTAAAGTTTA 4474
Db 611 GCGTGTATTAATCTTACGTTCACTGTGTGATTTTGAAGTTTGGTAGCTGTGGTT 670
4475 GTTGGTATTTATATATGATGTTTATTTGTTAGTTTGGATCGTTTGTGTT 4534
Qy 671 TCTTTGG 677
4535 TTTTGG 4541
Db

RESULT 9
US-10-424-599-109362
Sequence 109362, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
PRIOR FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 109362
LENGTH: 712
TYPE: DNA
ORGANISM: Glycine max

FEATURE:
NAME/KEY: unsure
LOCATION: (1) (712)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_69769C.1
US-10-424-599-109362

Query Match
Best Local Similarity 9.3%; Score 64.6; DB 12; Length 712;
Matches 169; Conservative 0; Mismatches 175; Indels 0; Gaps 0;

Db
328 TGTCTCGATTCTCTTTCTTCGCTGGAATAATGCTATGTTCTGATTTGGAAG 387
173 TGTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 232
Qy 388 GTTTTGTCTATGAGTTACTTTTCCCTATATTATATGTTCTTAGTAAAGTAACT 447
233 TTTTCTTCTGCTCTTTGTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 292
Qy 448 GCGCTTACTGTTTGTGATTTGTTGCTTCACTGCTTCACTGCTGATCGAGTA 507
293 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 352
Qy 508 TTTGACTGTGAAATCCTCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGT 567
353 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 412
Qy 568 CTTTGTGCTTGTATGTTTGTGTTTGTGAGCTATGCTGTTGCTTGTGTTTGTGTT 627
413 TTTTGTGCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 472
Qy 628 CGTCAATGNGGATTTGAGATTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 671
473 TTTTGGGGGGGCTGCTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTTTGTGTT 516
Db

RESULT 10
US-10-311-455-1669
Sequence 1669, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311,455
PRIOR FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 1669
LENGTH: 6668
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
NAME/KEY: unsure
LOCATION: 4733
OTHER INFORMATION: n is a or g or c or t
US-10-311-455-1669

Query Match
Best Local Similarity 9.3%; Score 64.6; DB 14; Length 6668;
Matches 169; Conservative 0; Mismatches 174; Indels 0; Gaps 0;

QY 321 TGCATCTGCTGCATATTCCTTTCTGCTGAAAAATGCCCTAATGCTGCAT 380
 Db 5264 TGAGTTCGGTTTGAATTTTTTTTTTTTGAAGATGTTGTTTTTTTTT 5323
 QY 381 TTGGAAGTTTGTGCTAGGAGTTACTTTTCCCTAATTTATAGTTCTTAGTAC 440
 Db 5324 TCGTTCGTT 5383
 QY 441 GAACTCGCTCTACTGTTTTTTTGTCAATTTGTTGCTTACCGTTAGTGCAT 500
 Db 5384 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTATTTGGAATGTTTGAGATT 5443
 QY 501 CGAGATTATGACTGGAATAATCCCTGTTTTTGTGTTTCATATAATCGAT 560
 Db 5444 TTTTTTTTTTTTTTTGAAGTTGTGTTTTTTTTTTTTTTCGTTTTTTTTTT 5503
 QY 561 TGAATCACTTTGCTCTGATGATTTGTTTTTGAAGCTATGCTTGTGCTTAT 620
 Db 5504 TT 5563
 QY 621 AACTTCACTTCACTGCTGATTTTGAATTTTGGATGACT 663
 Db 5564 TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTAT 5606

RESULT 11
 US-10-311-455-1692
 ; Sequence 1692, Application US/10311455
 ; Publication No. US20030143606A1
 ; GENERAL INFORMATION:
 ; APPLICANT: OLEK, Alexander
 ; APPLICANT: PISSENBROCK, Christian
 ; APPLICANT: BERLIN, Kurt
 ; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Detect
 ; TITLE OF INVENTION: cytosine methylation
 ; FILE REFERENCE: 5013.1014
 ; CURRENT APPLICATION NUMBER: US/10/311.455
 ; CURRENT FILING DATE: 2002-12-16
 ; PRIOR APPLICATION NUMBER: PCT/EP01/07537
 ; PRIOR FILING DATE: 2001-07-02
 ; PRIOR APPLICATION NUMBER: DE 10032529.7
 ; PRIOR FILING DATE: 2000-06-30
 ; PRIOR APPLICATION NUMBER: DE 10043826.1
 ; PRIOR FILING DATE: 2000-09-01
 ; NUMBER OF SEQ ID NOS: 2424
 ; SEQ ID NO 1692
 ; LENGTH: 17934
 ; TYPE: DNA
 ; ORGANISM: Artificial Sequence
 ; FEATURE:
 ; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
 US-10-311-455-1692

Query Match 9.3%; Score 64.4; DB 14; Length 17934;
 Best Local Similarity 48.9%; Pred. No. 0.00014;
 Matches 173; Conservative 0; Mismatches 181; Indels 0; Gaps 0;
 QY 325 TTCTGCTGCATTTCTTTCTTCTGCGAAAAATGCCCTAATGCTGCATTCG 384
 Db 17579 TTGG 17638
 QY 385 AAGCTTTTGGCAGGAGTTACTTTTCCCTAATTTATAGTTCTTAGACAGATA 444
 Db 17639 TTTTTTTTTTTTTTTTATTTTTTTTTTTTTTTTTTTTTTTTTTTT 17698
 QY 445 CCGCGCTTACTGTTTGTTCATTTGTTGCTTTCACCGTTAGTGCATCGGA 504
 Db 17699 TT 17758
 QY 505 GATTTGACTGGAATAATCCCTGTTTTTGTGTTGTTGTTGTTGTTGTTGTTG 564
 Db 17759 TTTTTTTTTTTTTTTAGTTTTTTTTTTTTTTTTTTTTTTTATTTTTTTTTTT 17818

QY 565 CTACCTTTGCTGATGATGTTTTTTTGAAGCTAATGCGTTGCTGTTAATCT 624
 Db 17819 TTTTTTTTTTTTTTTGATTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 17878
 QY 625 TCACCTCATGCTGATGATTTTGAATTTTGGTACTGCTGCTTCTTGGT 678
 Db 17879 TTTTTTTTTTTTTTTTATTTGTTTTTTGTTTGAATGAGGATTAATAGT 17932

RESULT 12
 US-10-424-599-102083/c
 ; Sequence 102083, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa, Thomas J
 ; APPLICANT: Kovalic, David K
 ; APPLICANT: Zhou, Yihua
 ; APPLICANT: Cao, Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
 ; FILE REFERENCE: 38-21(53223)B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 102083
 ; LENGTH: 1214
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(1214)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_63196C.1
 US-10-424-599-102083

Query Match 9.1%; Score 63.4; DB 12; Length 1214;
 Best Local Similarity 46.9%; Pred. No. 6.5e-05;
 Matches 157; Conservative 0; Mismatches 178; Indels 0; Gaps 0;
 QY 325 TTCTGCTGCATTTCTTTCTTCTGCGAAAAATGCCCTAATGCTGCATTCG 384
 Db 1169 TTTTTTNNNTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT 1110
 QY 385 AAGCTTTTGGCAGGAGTTACTTTTCCCTAATTTATAGTTCTTAGACAGATA 444
 Db 1109 NTNTTT 1050
 QY 445 CCGCGCTTACTGTTTGTTCATTTGTTGCTTTCACCGTTAGTGCATCGGA 504
 Db 1049 TTNTTTTTTT 990
 QY 505 GATTTGACTGGAATAATCCCTGTTTTTGTGTTGTTGTTGTTGTTGTTGTTG 930
 Db 989 TTATTTTTTTT 930
 QY 565 CTACCTTTGCTGATGATGTTTGTGTTTGAAGCTAATGCTGCTGTTAATCT 624
 Db 929 TTTCTTTTTTTATTTTTTTTTTTTTTTTTTTTATTTATTTATTTTNTTTTT 870
 QY 625 TCACCTCATGCTGATGATTTGAGATTTGGTAGT 659
 Db 869 TTTTTTTATTTNNNTTCTTTTTTTTTTTTTTTT 835

RESULT 13
 US-09-983-965-2109
 ; Sequence 2109, Application US/09983965
 ; Patent No. US20020137160A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Warren, Wesley C.
 ; APPLICANT: Tao, Nenping
 ; APPLICANT: Byatt, John C.

APPLICANT: Mathalaagan, Nagappan
TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
FILE REFERENCE: 37-21(10297)C
CURRENT APPLICATION NUMBER: US/09/983,965
CURRENT FILING DATE: 2001-10-26
PRIOR APPLICATION NUMBER: US 09/465,231
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: US 60/113,678
PRIOR FILING DATE: 1998-12-17
NUMBER OF SEQ ID NOS: 5912
SEQ ID NO 2109
LENGTH: 529
TYPE: DNA
ORGANISM: Bos taurus
FEATURE:
NAME/KEY: unsure
LOCATION: (335)
OTHER INFORMATION:
US-09-983-965-2109

Query Match 9.0%; Score 62.8; DB 9; Length 529;
Best Local Similarity 49.0%; Pred. No. 6e-05;
Matches 166; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 337 TTTCTCTTTTCTGCGTGAAGAAATTCGCCCTATGTCGATTCGAAAGTTTGTG 396
DB 53 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 112
QY 397 CTATGGTACTTCTTCCCTATTTTATGTTCTTAGTAAAGATTCGCGCTTAC 456
DB 113 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 172
QY 457 TGTCTTGTCTATTTGTGCTTTCACCGTTTACGCGTAAAGATTTGACGT 516
DB 173 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 232
QY 517 GAAATACCTGCTTTTGTGTTTCTTATATTAACGATTCACCTTTGTG 576
DB 233 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 292
QY 577 CTATGATGTTGTTTGTAGCCTATGCGTTGTGCTGTATTAACCTCACGTCATG 636
DB 293 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 352
QY 637 GTGATTTGAGATTTGTGAGTGTGCTGTGCTTTCTT 675
DB 353 TTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTT 391

RESULT 14
US-10-125-968-78/c
Sequence 78, Application US/10125968
Publication No. US20030215805A1
GENERAL INFORMATION:
APPLICANT: Illie, James
APPLICANT: Palermo, Adam
APPLICANT: Wang, Youzhen
APPLICANT: Steinmann, Kathleen
APPLICANT: Elias, Josh
APPLICANT: Mertens, Maureen
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND
TITLE OF INVENTION: METHODS FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
FILE REFERENCE: MRI-032
CURRENT APPLICATION NUMBER: US/10/125,968
CURRENT FILING DATE: 2002-04-19
PRIOR APPLICATION NUMBER: US 60/285,163
PRIOR FILING DATE: 2001-04-20
NUMBER OF SEQ ID NOS: 1417
SOFTWARE: FastSeq for windows Version 4.0
SEQ ID NO 78

LENGTH: 629
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: 25, 63, 64, 65, 142, 158, 159, 160, 204, 223, 224, 233, 255,
LOCATION: 256, 257, 258, 260, 263, 270, 271, 272, 286, 287, 290, 291,
LOCATION: 292, 293, 295, 296, 297, 298, 299, 303, 324, 333, 349,
LOCATION: 351, 352, 353, 354, 355, 356, 357, 358, 359, 360
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc.feature
LOCATION: 361, 362, 363, 364, 365, 366, 367, 369, 370, 371, 372,*373,
LOCATION: 374, 375, 376, 377, 380, 384, 386, 389, 390, 391, 393, 406,
LOCATION: 408, 409, 417, 419, 421, 424, 429, 434, 438, 439, 451, 453,
LOCATION: 456, 474, 475, 484, 486, 489, 492, 493, 501, 505, 516
OTHER INFORMATION: n = A,T,C or G
FEATURE:
NAME/KEY: misc.feature
LOCATION: 523, 536, 537, 538, 543, 546, 547, 554, 555, 558, 563, 565,
LOCATION: 568, 571, 572, 574, 575, 577, 581, 582, 586, 587, 592, 594,
LOCATION: 610, 611, 629
OTHER INFORMATION: n = A,T,C or G
US-10-125-968-78

Query Match 9.0%; Score 62.8; DB 15; Length 629;
Best Local Similarity 39.7%; Pred. No. 6.5e-05;
Matches 121; Conservative 0; Mismatches 184; Indels 0; Gaps 0;

QY 371 TGTCTCGATTTGGAAGTTTGTGCTATGCGTACTTTTCCCTATATTTATAGT 430
DB 617 TTTTNNNTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTTCTTTT 558
QY 431 CTAGGTAAAGTACTGCGTCTTACTGTTTGTTCATTTGTGTGCTTACCGTT 490
DB 557 TTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTT 498
QY 491 AGTGGCTGATCGAGATTTGACTGGAATAACCTGTTTGTGTTGTTTCATA 550
DB 497 TTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTT 438
QY 551 TAAATGATGATCTACCTTTGTGCTTGAATGTTTGTGAGCCTATGCGTTG 610
DB 437 TTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTTTNNNTTT 378
QY 611 GCGTGTATTAATCTACGTTCAATGTCGATTTTGAATTTTGTAGTACGTCG 670
DB 377 NNNNNNNNTNN 318
QY 671 TCTTT 675
DB 317 TTTT 313

RESULT 15
US-10-221-714A-502
Sequence 502, Application US/10221714A
Publication No. US20040048254A1
GENERAL INFORMATION:
APPLICANT: Olek, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with
TITLE OF INVENTION: tumor suppressor genes and oncogenes
FILE REFERENCE: 5013.1005
CURRENT APPLICATION NUMBER: US/10/221,714A
CURRENT FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: PCT/EP01/02955
PRIOR FILING DATE: 2001-03-15
PRIOR APPLICATION NUMBER: DE 10013847.0
PRIOR FILING DATE: 2000-03-15
PRIOR APPLICATION NUMBER: DE 10019058.8
PRIOR FILING DATE: 2000-04-06

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using bw model

Run on: April 7, 2004, 14:30:53 ; Search time 4901.81 Seconds

(without alignments)
16605.743 Million cell updates/sec

Title: US-10-030-829-2

Perfect score: 1878
Sequence: 1 atgagcttcagagctgcgcgc.....ttcacatgagatgatga 1878

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 2167151695 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

```

1: gb_ba:*
2: gb_hcg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_scs:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_scs:*
28: em_un:*
29: em_vl:*
30: em_hcg_hum:*
31: em_hcg_inv:*
32: em_hcg_other:*
33: em_hcg_mus:*
34: em_hcg_pin:*
35: em_hcg_rod:*
36: em_hcg_mam:*
37: em_hcg_vrt:*
38: em_sy:*
39: em_hcgo_hum:*
40: em_hcgo_mus:*
41: em_hcgo_other:*

```

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1878	100.0	1878	6	AX078761 Sequence
2	1878	100.0	1878	6	BT004380 Arabidops
3	1878	100.0	2162	8	BT002944 Arabidops
4	1462	77.8	2254	8	AF239719 Arabidops
5	1462	77.8	3275	6	AX078760 Sequence
6	1462	77.8	8165	8	AB025633 Arabidops
7	437.8	23.3	650	8	AT528171 Arabidops
8	427.4	22.8	2181	8	AK064995 Oryza sat
9	372	19.8	2570	8	AK100699 Oryza sat
10	368	19.6	1556	8	AK064217 Oryza sat
11	291.6	15.5	1564	8	AF542974 Trilicium
12	288.8	15.4	1214	8	AK061975 Oryza sat
13	192	10.2	513	8	AF69493 Trilicium
14	169.4	9.0	91053	8	CNS08CAO Arabidops
15	169.4	9.0	112231	8	CNS07Y01 Arabidops
16	130.4	6.9	94911	8	AT114E10 Arabidops
17	88.4	4.7	179685	10	AC122056 Arabidops
18	87.8	4.6	204635	10	AL928594 Arabidops
19	86.6	4.5	282861	2	AC111701 Arabidops
20	84	4.4	227272	2	AC103478 Arabidops
21	83.4	4.4	279587	2	AC112408 Arabidops
22	83.2	4.4	157393	10	AC124199 Arabidops
23	82.4	4.3	248672	2	AC105804 Arabidops
24	80.2	4.2	124244	10	AL929026 Arabidops
25	79.4	4.2	188782	10	AL713870 Mouse DNA
26	79.4	4.2	195630	10	AC098726 Mus muscu
27	79	4.2	114848	9	AC018664 Homo sapi
28	78.6	4.2	178635	9	AC008080 Homo sapi
29	78.6	4.2	194006	10	AL732441 Arabidops
30	78.6	4.2	233713	2	AC133034 Arabidops
31	78	4.2	237549	2	AC112451 Arabidops
32	78	4.1	201074	2	AC102097 Mus muscu
33	77.8	4.1	211947	10	AC107828 Arabidops
34	77.8	4.1	188587	10	AC124464 Arabidops
35	77.6	4.1	205037	10	AC126266 Arabidops
36	77.6	4.1	169256	10	AC127264 Arabidops
37	77.4	4.1	178927	2	AC118897 Arabidops
38	77.4	4.1	217511	10	AC125357 Arabidops
39	77.4	4.1	56052	8	AP006363 Lotus cor
40	77.2	4.1	215910	10	AC127335 Arabidops
41	77	4.1	195987	10	AC122537 Arabidops
42	76.8	4.1	197599	2	AC119769 Arabidops
43	76.8	4.1	226601	10	AC122863 Arabidops
44	76.8	4.1	248718	10	AC117629 Arabidops
45	76.6	4.1	248718	10	AC117629 Arabidops

ALIGNMENTS

```

RESULT 1
AX078761 1878 bp DNA 1linear PAT 22-FEB-2001
LOCUS AX078761
DEFINITION Sequence 2 from Patent WO0105951.
ACCESSION AX078761
VERSION AX078761.1 GI:13158380
KEYWORDS
SOURCE
ORGANISM
Arabidopsis thaliana (thale cress)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE
1. Beclin,C., Elmayer,T. and Vaucheret,H.
AUTHORS
Novel gsg3 plant gene and use thereof
TITLE

```

JOURNAL

Patent: WO 0105951-A 2 25-JAN-2001;
 AVENTIS CROPS SCIENCE S.A.. (FR) ; INSTITUT NATIONAL DE LA RECHERCHE
 AGRONOMIQUE (FR)

FEATURES

Location/Qualifiers
 1..1878

/organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

CDS

1..1878
 /code="unassigned protein product"

/protein_id="CAC32419.1"
 /db_xref="GI:13158381"

/translation="MSSRAGPMSKEKNGVGVYRPEVQLVGLAETRLASQDGGEW
 EYISKNNKPKGTSKTVSONSPRAMGQOQGRGNSGGRNYSGRNGRG
 IOANISGRGRALSRKYDNPFVA PYPVSRPLRGLGMMWMOAGSAOHTAVQEPVEDD
 VDAAREEDSDALDSDDDLDVDPVDSQKSHSGRKNMFKPKFSLSLIEO
 INEPOQWHCPACQNGPGALDWNHPLAHPKARVYKRLRELAEVLEEDLDMRG
 ASVIPCGLIYQWKGIGDEKDEYIVPMPVIMNTRLDKDNKMLGNOBELSYF
 DKREALRARSYGPQGRGMSVLMESATGYLEARLRELAEMLGDRILAWQKRM
 FSGVRYLYGFLATKODLDIFNHSQKTRLEFELSVOEMVVEKELROI SEDNOOLNY
 FNKJL SKONKHA KVLSESL EIMSEKLR TAEADNR IYRSTKQHEQNEEMDAHDF
 MSIKRIHERBDAKENPFMLQOQERAKVVGQOQOONINSNDPCR KRAEVSFIEF
 QEKEMEFEEVEREMLIKDEKMEKMKRHHHEIFDLEREPDLAEQLMYRKLHNEED
 D"

ORIGIN

Query Match 100.0%; Score 1878; DB 6; Length 1878;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 1878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTTCTAGGCTGGTCCAAATGCTTAAGAAAAGAAAGCTTCAGGCTGTTATAGCCCT 60
 DB 1 ATGAGTTCTAGGCTGGTCCAAATGCTTAAGAAAAGAAAGCTTCAGGCTGTTATAGCCCT 60
 QY 61 GAGGTTGAAGATTGTTCAAGTTGGCAGGAGCAGACTGGCTTCTTCAAGATGAT 120
 DB 61 GAGGTTGAAGATTGTTCAAGTTGGCAGGAGCAGACTGGCTTCTTCAAGATGAT 120
 QY 121 GAGAGAGATGGAGGTCTATTTCCAGAGAAACAGAAACCAAGCAACTTCTGGA 180
 DB 121 GAGAGAGATGGAGGTCTATTTCCAGAGAAACAGAAACCAAGCAACTTCTGGA 180
 QY 181 AAAAATTGGGTTTCTCAAAATCTCTTCAAGCTTGGGTTGTCAGCAAGGG 240
 DB 181 AAAAATTGGGTTTCTCAAAATCTCTTCAAGCTTGGGTTGTCAGCAAGGG 240
 QY 241 AGAGGTAGCAACGATCTGGGAGAGAAACATGATCTCGGAGAGGTAACGCAATGCT 300
 DB 241 AGAGGTAGCAACGATCTGGGAGAGAAACATGATCTCGGAGAGGTAACGCAATGCT 300
 QY 301 CGGGGCAATTCAGCTAATCATATCTGTCGGGAGCAGACGTTGAGCAAAAGTATGATAAC 360
 DB 301 CGGGGCAATTCAGCTAATCATATCTGTCGGGAGCAGACGTTGAGCAAAAGTATGATAAC 360
 QY 361 AACTTTGTGCAACCCCACTGTATCTCGCCCTCTTTGGAAGAGATGGAATTGGCAG 420
 DB 361 AACTTTGTGCAACCCCACTGTATCTCGCCCTCTTTGGAAGAGATGGAATTGGCAG 420
 QY 421 GCAAGAGAGGTTCTGCTCAGACACAGCTGTCAGAGAGTTCTCTGACGTGAGAGATGAT 480
 DB 421 GCAAGAGAGGTTCTGCTCAGACACAGCTGTCAGAGAGTTCTCTGACGTGAGAGATGAT 480
 QY 481 GTGATATATGCTTCTGAGAGAGAGATGATTCGATGCTTTGATGATTCGATGACGAC 540
 DB 481 GTGATATATGCTTCTGAGAGAGAGATGATTCGATGCTTTGATGATTCGATGACGAC 540
 QY 541 CTTCGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 DB 541 CTTCGCAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 QY 601 AATAAGTGTCTCAAAAAGTTCTTTGGCAGCTTGTGATGATGATGATGATGATGATGATGAT 660

DB 601 AATAAGTGTCTCAAAAAGTTCTTTGGCAGCTTGTGATGATGATGATGATGATGATGATGAT 660
 QY 661 GAACCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 DB 661 GAACCAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
 QY 721 TATTAACCTGCAACCTCTTACTACTAGCTCATGCGAGCAAAAGAGATGACGAGTTAAGCTC 780
 DB 721 TATTAACCTGCAACCTCTTACTACTAGCTCATGCGAGCAAAAGAGATGACGAGTTAAGCTC 780
 QY 781 CATAGAGATTTGCTGAAATTTTGAAGAGATCTACAGATGAGAGCGGCTCTGAT 840
 DB 781 CATAGAGATTTGCTGAAATTTTGAAGAGATCTACAGATGAGAGCGGCTCTGAT 840
 QY 841 CTTGTGTGATGATTTATGGCAGTGGAGGTTTGGGTGAGATGAAAGATTAAGAA 900
 DB 841 CTTGTGTGATGATTTATGGCAGTGGAGGTTTGGGTGAGATGAAAGATTAAGAA 900
 QY 901 ATTGTGTGCTCCAAATGCTGATCATGTAATCTAGACTGATTAAGAGATTAAGAT 960
 DB 901 ATTGTGTGCTCCAAATGCTGATCATGTAATCTAGACTGATTAAGAGATTAAGAT 960
 QY 961 AAGTGGCTCGGATGAGGCAACCAAGAGCTGCTGAATTCCTGCAAGATGAGGCTTT 1020
 DB 961 AAGTGGCTCGGATGAGGCAACCAAGAGCTGCTGAATTCCTGCAAGATGAGGCTTT 1020
 QY 1021 AGAGCAAGCCTATTCCTATGCTCCACAGAGCCATGCTGGAGATGAGATGATGATGAT 1080
 DB 1021 AGAGCAAGCCTATTCCTATGCTCCACAGAGCCATGCTGGAGATGAGATGATGATGAT 1080
 QY 1081 AGCAGTGCACATGCTATTTGAGAGCCCAAGCCTCCACCGGAGATTAGCTGAGATGAGG 1140
 DB 1081 AGCAGTGCACATGCTATTTGAGAGCCCAAGCCTCCACCGGAGATTAGCTGAGATGAGG 1140
 QY 1141 TTAGATAGAAATGCTGAGGCTGAGAGGCGAGATGTTTCTGAGAGTGTGCTCAATG 1200
 DB 1141 TTAGATAGAAATGCTGAGGCTGAGAGGCGAGATGTTTCTGAGAGTGTGCTCAATG 1200
 QY 1201 TATGCTTCTTGAACAGAGAGAGATCTGAGATCAATCAACATCTCAAGGCAAA 1260
 DB 1201 TATGCTTCTTGAACAGAGAGAGATCTGAGATCAATCAACATCTCAAGGCAAA 1260
 QY 1261 ACAAGGCTGAAATTTGAGTTGAATCATACAGAGATGTTTGAAGAGAGCTGAGGAG 1320
 DB 1261 ACAAGGCTGAAATTTGAGTTGAATCATACAGAGATGTTTGAAGAGAGCTGAGGAG 1320
 QY 1321 ATCTCTGAGAGCAATCAGAGCTGAACTTAAAGAACTTCAAAACAGAAACAG 1380
 DB 1321 ATCTCTGAGAGCAATCAGAGCTGAACTTAAAGAACTTCAAAACAGAAACAG 1380
 QY 1381 CAGGCAAGGCTTGAAGATCTCTGAAATTTATGAGAGAGAGCTGCTGAACTGCA 1440
 DB 1381 CAGGCAAGGCTTGAAGATCTCTGAAATTTATGAGAGAGAGCTGCTGAACTGCA 1440
 QY 1441 GAGGATATCGATCGATGAG 1500
 DB 1441 GAGGATATCGATCGATGAG 1500
 QY 1501 ATGATGACACAGCAGGTTTTCATGATTCATTAACAACAATCCATTAAGAGAGAGAG 1560
 DB 1501 ATGATGACACAGCAGGTTTTCATGATTCATTAACAACAATCCATTAAGAGAGAGAG 1560
 QY 1561 GCAAGAGAGAGAGATTTGAGAGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 DB 1561 GCAAGAGAGAGAGATTTGAGAGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 QY 1621 CAGCAGCAGAACTTAATCTCTTACAAATGAGATGCTGAGAGAGAGAGAGAGAGAGAGAG 1680
 DB 1621 CAGCAGCAGAACTTAATCTCTTACAAATGAGATGCTGAGAGAGAGAGAGAGAGAGAGAG 1680
 QY 1681 TCAAGCTTATGAGATTTCAAG 1740

Db 1681 TCAAGCTTCATCGAGTTTCAAGAGAAAGAGATGAGAGAGTTGTGAGAGAGAGGAGATG 1740

Qy 1741 CTGATAAAGATCAAGAGAAAGATGAGAGATGAGAGAGGATCGAGAGATG 1800

Db 1741 CTGATAAAGATCAAGAGAAAGATGAGAGATGAGAGAGGATCGAGAGATG 1800

Qy 1801 TTGTATCTGGAGAGAAAGATTTGATGAGGCTTTTGAAACAGCTCATGTACAGATGAGCTT 1860

Db 1801 TTGTATCTGGAGAGAAAGATTTGATGAGGCTTTTGAAACAGCTCATGTACAGATGAGCTT 1860

Qy 1861 CACATGAAAGATGATGCA 1878

Db 1861 CACATGAAAGATGATGCA 1878

RESULT 2

BT004380 1909 bp mRNA linear PLN 14-FEB-2003

LOCUS Arabidopsis thaliana clone U20243 unknown protein (At5g23570) mRNA, complete cds.

ACCESSION BT004380

VERSION BT004380.1 GI:28393932

KEYWORDS FLI_CDN.

SOURCE Arabidopsis thaliana (thale cress)

ORGANISM Arabidopsis thaliana

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustoids II; Brassicales; Brassicaceae; Arabidopsis.

REFERENCE 1 (bases 1 to 1909)

AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W., Becker, J.R. and Theologis, A.

TITLE Arabidopsis Open Reading Frame (ORF) Clones

JOURNAL Unpublished

REFERENCE 2 (bases 1 to 1909)

AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Hsuan, V.W., Lee, J.M., Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y., Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M., Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P., Southwick, A., Tripp, M.G., Wu, T., Shinzaki, K., Davis, R.W., Becker, J.R. and Theologis, A.

TITLE Direct Submission

JOURNAL Submitted (14-FEB-2003) Plant Gene Expression Center, 800 Buchanan Street, Albany, CA 94710, USA

COMMENT The RIKEN Genomic Sciences Center (GSC) members carried out the collection and clustering of RAPL cDNAs (RAPL cDNA: RIKEN Arabidopsis Full-Length cDNA): Seki, M., Narusaka, M., Ishida, J., Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J., Hayashizaki, Y. and Shinzaki, K.

FEATURES

Source

Location/Qualifiers

1..1909

/organism="Arabidopsis thaliana"

/mol_type="mRNA"

/db_xref="taxon:3702"

/chromosome="5"

/clone="U20243"

/ecotype="Columbia"

/note="this clone is in punit 51."

1..1909

/gene="At5g23570"

1..1878

/gene="At5g23570"

/codon_start=1

/evidence="experimental"

/product="unknown protein"

/protein_id="AA042374.1"

/db_xref="GI:28393932"

/translation="MSSRAGPWSKREKRVGGYRPERVEOLVGLAGTRLLASODDGESEVVISKKNRKPENITSGKTIVSGNSNPAPRAGGQOQGRGNSVGRGNVSGRGNNGRGIOANISGRGRLSRKTDNNFVAPPPVSRPLSGMNMQARGSAQHTAVQEPFVEDDVNNAEENDSDALDDSDDLASDDYDSVSGSHSGSRKQNMFKFQFSLDLSLBQINBPORHWPACQNGPGLADWYNLPLAHLARTGARVYKLHRELAETLELDLQWGLASVIPCGETIYQGMKGLGEDEKDYELIWPMPVILNMTRLDKDNDKMLGNGBLEYFDKVALRARHSYGPQHGMSVLMFESSATGYLEAERLYHRELAEMGLDIRIANGQKSMFSGYRQLYGLATTKQDDLPVHSGQGRRLKFLKSYQYEMVKELRQISENQOQNTFNNKLSKQNTAKVLEESLEINSEKLRRTAEDNRIVRQRTKQHEONREEMADHDEFMDSIKQIHERRDPAKENEEMLQOQERARVVGQOQONINSSNDCKRKAEVVSFLEFDEKEMBEFVEREMFLIKQEKMEEMKRAHHEIFDLEREFDLEALQMLMYKQLHNEFD"

1879..1909

/gene="At5g23570"

ORIGIN

3'UTR

Query Match 100.0%; Score 1878; DB 8; Length 1909;

Best local Similarity 100.0%; Pred. No. 0;

Matches 1878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGAGTTCTAGAGGCTGCTGTCATATGTTCAAGAAAGACGTTCAAGGCTGTTATGAGCCT 60

Db 1 ATGAGTTCTAGAGGCTGCTGTCATATGTTCAAGAAAGACGTTCAAGGCTGTTATGAGCCT 60

Qy 61 GAGGTTGAACATGTTGTTCAAGGTTGGAGGAGGAGCTGCTTCAAGAATGAT 120

Db 61 GAGGTTGAACATGTTGTTCAAGGTTGGAGGAGGAGCTGCTTCAAGAATGAT 120

Qy 121 GGAGGAGAGTGGAGAGCTATTTCCAAAGAAACAAGAACCAAGAAACACTTCTGGA 180

Db 121 GGAGGAGAGTGGAGAGCTATTTCCAAAGAAACAAGAACCAAGAAACACTTCTGGA 180

Qy 181 AAAACTGGGTTTCTCAGAAATTCCTCTTAAGACTTGGGCTGTCAGACGAGG 240

Db 181 AAAACTGGGTTTCTCAGAAATTCCTCTTAAGACTTGGGCTGTCAGACGAGG 240

Qy 241 AAGGTAAGCAAGTATCTGGGAGAGAAACAATGATTCGGGAGAGTTAACCGCAATGCT 300

Db 241 AAGGTAAGCAAGTATCTGGGAGAGAAACAATGATTCGGGAGAGTTAACCGCAATGCT 300

Qy 301 CGGGGCAATCAAGCTAACATATCTGCTCGGGGACGAGCGTTGAGCGAAAGATGATAC 360

Db 301 CGGGGCAATCAAGCTAACATATCTGCTCGGGGACGAGCGTTGAGCGAAAGATGATAC 360

Qy 361 AACTTGTGGCAACCCCACTTATCTCGCCCTCTTTGGAAGAGAGATGAAATTTGGCAG 420

Db 361 AACTTGTGGCAACCCCACTTATCTCGCCCTCTTTGGAAGAGAGATGAAATTTGGCAG 420

Qy 421 GCAAGAGAGGTTTCTGCTCAGACACAGCTGTGCGAGATTTCTTGAAGATGATAT 480

Db 421 GCAAGAGAGGTTTCTGCTCAGACACAGCTGTGCGAGATTTCTTGAAGATGATAT 480

Qy 481 GTGATTAATGCTTCTGAGAGAGAGATGTTCCGATGCTTTGAGATGATTTCTGATGCGAC 540

Db 481 GTGATTAATGCTTCTGAGAGAGAGATGTTCCGATGCTTTGAGATGATTTCTGATGCGAC 540

Qy 541 CTTGCAAGTATGATTAATGATCTCGAGATGATGATCAAAAAGCCATGATCAAGAAAGCAG 600

Db 541 CTTGCAAGTATGATTAATGATCTCGAGATGATGATCAAAAAGCCATGATCAAGAAAGCAG 600

QY 601 AATAAGTGTTCAAAAAGTTCTTTGGCAGCTTGGAATGCTTGTGATGAGCAGATAAT 660
 DB 601 AATAAGTGTTCAAAAAGTTCTTTGGCAGCTTGGAATGCTTGTGATGAGCAGATAAT 660
 QY 661 GAACCAAGAGGAGGAGTGTCCAGCTTGTGAGAACGAGCCTGTGCTCATGATGG 720
 DB 661 GAACCAAGAGGAGGAGTGTCCAGCTTGTGAGAACGAGCCTGTGCTCATGATGG 720
 QY 721 TATAAACCCTGACCTCTTACTGATGAGAGACAAAAGAGCTTAGGCGAGTTAACTC 780
 DB 721 TATAAACCCTGACCTCTTACTGATGAGAGACAAAAGAGCTTAGGCGAGTTAACTC 780
 QY 781 CATAGAGAAATTTGGCTGAATTTTAAAGAAAGATCTACAGATGAGAGGCGCATCTGCAT 840
 DB 781 CATAGAGAAATTTGGCTGAATTTTAAAGAAAGATCTACAGATGAGAGGCGCATCTGCAT 840
 QY 841 CCTTGTGTGAGATTTATGTGGCAGTGGAGAGGTTTGGTGAAGATGAAAAGATTAATGA 900
 DB 841 CCTTGTGTGAGATTTATGTGGCAGTGGAGAGGTTTGGTGAAGATGAAAAGATTAATGA 900
 QY 901 AATTGTGCTGCTTCAATGCTCATCATGATTAATCTAGACTGATTAAGACATTAACGAT 960
 DB 901 AATTGTGCTGCTTCAATGCTCATCATGATTAATCTAGACTGATTAAGACATTAACGAT 960
 QY 961 AAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 DB 961 AAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1020
 QY 1021 AGAGCAAGCCTATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 DB 1021 AGAGCAAGCCTATCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1080
 QY 1081 AGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 DB 1081 AGCAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1140
 QY 1141 TTAGATAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 DB 1141 TTAGATAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1200
 QY 1201 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 DB 1201 TATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1260
 QY 1261 ACAAGGCTGAAATTCGATGAAATTCGATGAAATTCGATGAAATTCGATGAAATTCGAT 1320
 DB 1261 ACAAGGCTGAAATTCGATGAAATTCGATGAAATTCGATGAAATTCGATGAAATTCGAT 1320
 QY 1321 ATCTGAGAGCAATGAGCAGCTGAACTTAAAGAACAGCTCTCAAAACAGAACAG 1380
 DB 1321 ATCTGAGAGCAATGAGCAGCTGAACTTAAAGAACAGCTCTCTCAAAACAGAACAG 1380
 QY 1381 CACGCCAAGGTGCTTGAAGATCTCTGGAATTAATGAGCAGAAAGCTGCTGAAGCTGCA 1440
 DB 1381 CACGCCAAGGTGCTTGAAGATCTCTGGAATTAATGAGCAGAAAGCTGCTGAAGCTGCA 1440
 QY 1441 GAGATTAATGAGATCTGAG 1500
 DB 1441 GAGATTAATGAGATCTGAG 1500
 QY 1501 ATGAGATCAACACAGAGGTTTTCATGATGATCAATCAACAGATCCATGAAAGAGAGAC 1560
 DB 1501 ATGAGATCAACACAGAGGTTTTCATGATGATCAATCAACAGATCCATGAAAGAGAGAC 1560
 QY 1561 GCAAGAGAGAGAGATTTGAGATGTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 DB 1561 GCAAGAGAGAGAGATTTGAGATGTTGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1620
 QY 1621 CAGCAGAGAGAGATTTATCCCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680
 DB 1621 CAGCAGAGAGAGATTTATCCCTCTAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1680

QY 1681 TCAAGCTTCATCGAGTTTCAG 1740
 DB 1681 TCAAGCTTCATCGAGTTTCAG 1740
 QY 1741 CTGATTAAGATTCAGTA 1800
 DB 1741 CTGATTAAGATTCAGTA 1800
 QY 1801 TTTGATCTGAG 1860
 DB 1801 TTTGATCTGAG 1860
 QY 1861 CACATGAAGATGATGA 1878
 DB 1861 CACATGAAGATGATGA 1878
 RESULT 3
 BT002944 2162 bp mRNA linear PLAN 15-JAN-2003
 LOCUS Arabidopsis thaliana clone RAFL14-93-K05 (R20243) unknown protein
 DEFINITION (A5523570) mRNA, complete cds.
 ACCESSION BT002944 GI:27754622
 VERSION BT002944.1
 KEYWORDS FLI CDNA.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustroids II; Brassicales; Brassicaceae; Arabidopsis.
 REFERENCE
 AUTHORS Yamada, K., Chan, M.M., Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M.,
 Onodera, C.S., Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
 Yu, G., Yuan, S., Carninci, P., Chen, H., Cheuk, R., Hayashizaki, Y.,
 Ishida, J., Jones, T., Kamiya, A., Kawai, J., Kim, C.J., Narusaka, M.,
 Nguyen, M., Palm, C.J., Sakurai, T., Satou, M., Seki, M., Shim, P.,
 Southwick, A., Tripp, M.G., Wu, T., Shinozaki, K., Davis, R.W.,
 Ecker, J.R. and Theologis, A.
 Arabidopsis Full Length cDNA Clones
 Unpublished
 2 (bases 1 to 2162)
 TITLE
 JOURNAL
 REFERENCE
 AUTHORS
 COMMENT
 JOURNAL
 COMMENT
 The Salk, Stanford, PGEC (SSP) Consortium members carried out the
 sequencing and annotation of the RAFL cDNAs: Yamada, K., Chan, M.M.,
 Chang, C.H., Dale, J.M., Huan, V.W., Lee, J.M., Onodera, C.S.,
 Quach, H.L., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yu, G.,
 Yuan, S., Chen, H., Cheuk, R., Jones, T., Kim, C.J., Nguyen, M.,
 Palm, C.J., Shim, P., Southwick, A., Tripp, M.G., Wu, T., Davis, R.W.,
 Ecker, J.R. and Theologis, A.
 Yamada, K. (SSP/PGEC) and Seki, M. (RIKEN GSC) contributed equally to
 this work. Shinozaki, K. (RIKEN GSC) and Theologis, A. (SSP/PGEC)
 contributed equally to this work as PIs.
 Annotation based on July 2002 version of the Arabidopsis genome
 submitted to Genbank.
 FEATURES
 Location/Qualifiers
 1..2162
 source

```
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/db_xref="taxon:3702"
/chromosome="5"
/clone="RAF1.4-93-K05 (R20243)"
/ecotype="Columbia"
/notes="This clone is in a modified pBluescript vector2
(lambda PS) as a BamHI/XhoI insert."
1..2162
/gene="At5g23570"
1..91
/gene="At5g23570"
92..1969
/gene="At5g23570"
/codon_start=1
/evidence=experimental
/product="unknown protein"
/protein_id="AA022757.1"
/db_xref="GI:2754623"
/translation="MSSRAGPMSEKKNVQGYRPEVQLVQGLATRIASDDDGEM
EVISSKNKKPKGTSGKTWVSONSPNPRAGGQOQGRGSVSGRGNVSGRGNGRG
IQANISGRGRLSRKTDNPFVAPRPPVSRPLEGKMNQARSGSAHTAVDFEVED
VNDASEENDSDALDSDDDLSDYDSDVSQSHSRKQNKPKFPGSLDLSLEQ
INERQOMHCPCACONGPGLIDWNLHLPLAHARTKARVYKLELAELVLEKDLQMRG
ASVIPCSEIYGQKGLGEDEKQYEIVWPPVWVIMNTRLDKODNDKMLGMNDELLEYF
DKREALRARSYVGRPOGRHGVLMPESSATVYLEARLHRELAEMGLDRIAMGOKRM
FSGVROLYGLFLTKODDIFNHSOGSKTRKLELSYQEMVYKELRQISEDNQALY
FKNKSQKONKHAVALSESLINSKURTRADNRIVRQRTKQHEQUREMDNDHDF
MOSIKQIHERDRAKENFEMLDQOERAKVVGQOQONINPSNDCKRRAEVSFLIEF
OEKEMEFLVEBERMLIKDEKMKEMDKRHHEBIFLDEKEFDEALBOLMYKHGLAHND
D"
1970..2162
/gene="At5g23570"
misc_difference 2146
/gene="At5g23570"
/notes="not present in genomic sequence"

3' UTR

ORIGIN

Query Match 100.0%; Score 1878; DB 8; Length 2162;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGTTCTAGGCTGGTCCAAATGCTTAAGAAAGAAAGAAAGCTTCAGGCTGTTATAGGCTT 60
DB 92 ATGAGTTCTAGGCTGGTCCAAATGCTTAAGAAAGAAAGAAAGCTTCAGGCTGTTATAGGCTT 151
QY 61 GAGGTGAACAGTTGGTTCAAGGTTTGGCAGGAGCGAGACTGGCTTCTTCAAGATGAT 120
DB 152 GAGGTGAACAGTTGGTTCAAGGTTTGGCAGGAGCGAGACTGGCTTCTTCAAGATGAT 211
QY 121 GAGGAGAGTGGAGAGTCAATTTCCAGAGAAACAAGAAACAACACCTTCTTGA 180
DB 212 GAGGAGAGTGGAGAGTCAATTTCCAGAGAAACAAGAAACAACACCTTCTTGA 271
QY 181 AAAAATTGGGTTCTCGAATTCCTCTAGAGCTTGGGTTGTCAGCAGCAAGG 240
DB 272 AAAAATTGGGTTCTCGAATTCCTCTAGAGCTTGGGTTGTCAGCAGCAAGG 331
QY 241 AGAGTAGCAACGTATCTGGAGAGAGAAACAATGTATCCGAGAGAGTTAACGGCAATGCT 300
DB 332 AGAGTAGCAACGTATCTGGAGAGAGAAACAATGTATCCGAGAGAGTTAACGGCAATGCT 391
QY 301 CGGGGCAATTAACCTAATATCTGGTCCGGGAGCAGAGCTTGAACAGAAAGTATGATTAAC 360
DB 392 CGGGGCAATTAACCTAATATCTGGTCCGGGAGCAGAGCTTGAACAGAAAGTATGATTAAC 451
QY 361 AACTTTGGCAACCCCACTGTATCTCGCCCTCTTTGGAAGAGAGATGAATTTGGCAG 420
DB 452 AACTTTGGCAACCCCACTGTATCTCGCCCTCTTTGGAAGAGAGATGAATTTGGCAG 511
QY 421 GCAAGAGAGAGTTCTGTGACACACAGCTGTGCAAGAGATTTCTGACGTGAGAGATGAT 480
DB 512 GCAAGAGAGAGTTCTGTGACACACAGCTGTGCAAGAGATTTCTGACGTGAGAGATGAT 571
```

```
QY 481 GTGATTAATGCTTCTGAGGAAGAAATGATTCGATGCTTTGGATGATTTCTGATGACGAC 540
DB 572 GTGATTAATGCTTCTGAGGAAGAAATGATTCGATGCTTTGGATGATTTCTGATGACGAC 631
QY 541 CTTGCAAGTATGATTAATGATCTCGAGTGTGAGTCAAAAGACCATGATCAAGAAACGAC 600
DB 632 CTTGCAAGTATGATTAATGATCTCGAGTGTGAGTCAAAAGACCATGATCAAGAAACGAC 691
QY 601 AATAAGTGTTCAAAAGTCTTTGGCAGCTTGGATGATGCTGTGATGACGACGATTAAT 660
DB 692 AATAAGTGTTCAAAAGTCTTTGGCAGCTTGGATGATGCTGTGATGACGACGATTAAT 751
QY 661 GAACACAGAGGAGGAGGATGCTCCAGGCTTGCAGAAAGGACCTGATGATGATGATGATG 720
DB 752 GAACACAGAGGAGGAGGATGCTCCAGGCTTGCAGAAAGGACCTGATGATGATGATGATG 811
QY 721 TATTAACCTGACCCCTTACTAGCTCAATGCGAGGACAAAAGAGCTTAGGCGATTAAGCTT 780
DB 812 TATTAACCTGACCCCTTACTAGCTCAATGCGAGGACAAAAGAGCTTAGGCGATTAAGCTT 871
QY 781 CATAGGAATTTGGCTGAAGTTTGAAGAAAGATCTACATGAGAGAGCCGATCTGTCAAT 840
DB 872 CATAGGAATTTGGCTGAAGTTTGAAGAAAGATCTACATGAGAGAGCCGATCTGTCAAT 931
QY 841 CCTTGTGTGAGATTTATGGGAGTGGAAAGGTTTGGGTGAGGATGAAAAGATTAATGA 900
DB 932 CCTTGTGTGAGATTTATGGGAGTGGAAAGGTTTGGGTGAGGATGAAAAGATTAATGA 991
QY 901 ATTGTGTGCTCCCAATGATGATCATGAAATCTAGACTGATTAAGACATGAAT 960
DB 992 ATTGTGTGCTCCCAATGATGATCATGAAATCTAGACTGATTAAGACATGAAT 1051
QY 961 AATGTGCTGGCATTGGGCAACCAAGAGCTGTGGAATCTTGCACAAATATGAGGCTCTT 1020
DB 1052 AATGTGCTGGCATTGGGCAACCAAGAGCTGTGGAATCTTGCACAAATATGAGGCTCTT 1111
QY 1021 AGAGCAGGCAATTCCTATGATGATCAAGAGGCAATCGTGGAGTGAATGATGATGATG 1080
DB 1112 AGAGCAGGCAATTCCTATGATGATCAAGAGGCAATCGTGGAGTGAATGATGATGATG 1171
QY 1081 AGCAGTGCACCTGCTATTTTGAAGGCGCAACGCTTCAACCGGAGTTAGCTGATGAGG 1140
DB 1172 AGCAGTGCACCTGCTATTTTGAAGGCGCAACGCTTCAACCGGAGTTAGCTGATGAGG 1231
QY 1141 TTGATAGAAATTTGCTGGGCTCAAGACCGCAATGATTTTCTGAGAGTGTGCGCAACTG 1200
DB 1232 TTGATAGAAATTTGCTGGGCTCAAGACCGCAATGATTTTCTGAGAGTGTGCGCAACTG 1291
QY 1201 TATGCTTCTCTGCAACGAAGCAAGATCTGGAATCTAATCAATCAACTCTCAAGGCAAA 1260
DB 1292 TATGCTTCTCTGCAACGAAGCAAGATCTGGAATCTAATCAATCAACTCTCAAGGCAAA 1351
QY 1261 ACNAGGCTGAATTTGAGTTGAATATCATCAAGAGATGTTGTAAAGAGCTGAGGCAAG 1320
DB 1352 ACNAGGCTGAATTTGAGTTGAATATCATCAAGAGATGTTGTAAAGAGCTGAGGCAAG 1411
QY 1321 ATCTTGAAGACATAGAGAGCTGAATCTATTAAACAAGCTCTCAAAAACGAAACAAG 1380
DB 1412 ATCTTGAAGACATAGAGAGCTGAATCTATTAAACAAGCTCTCAAAAACGAAACAAG 1471
QY 1381 CAGGCCAAGTGTGTTAGGAATCTCTGGAATTTATAGAGGAGAAAGCTGGGTGAATGCA 1440
DB 1472 CAGGCCAAGTGTGTTAGGAATCTCTGGAATTTATAGAGGAGAAAGCTGGGTGAATGCA 1531
QY 1441 GAGGATTAATCGATCTGTGAGACAGAGAACTAAAGATCAGATCAAGACAGAGAAAG 1500
DB 1532 GAGGATTAATCGATCTGTGAGACAGAGAACTAAAGATCAGATCAAGACAGAGAAAG 1591
QY 1501 ATGATGCAACAGACAGGTTTTTTCATGATTTCAATTAACAGATCCATGAAGAAAGAGAC 1560
DB 1592 ATGATGCAACAGACAGGTTTTTTCATGATTTCAATTAACAGATCCATGAAGAAAGAGAC 1651
```

```

OY 1561 GCAAGAGAGAGATTTCAGATGTTTGCAGCAGCAGGAAAGCGCAAGGTTGTGGCAG 1620
DB 1652 GCAAGAGAGAGATTTCAGATGTTTGCAGCAGCAGGAAAGCGCAAGGTTGTGGCAG 1711
OY 1621 CAGCAGCAGAGATTTCAGATGTTTGCAGCAGCAGGAAAGCGCAAGGTTGTGGCAG 1680
DB 1712 CAGCAGCAGAGATTTCAGATGTTTGCAGCAGCAGGAAAGCGCAAGGTTGTGGCAG 1771
OY 1681 TCAAGCTTCATCAGATGTTTGCAGCAGCAGGAAAGCGCAAGGTTGTGGCAG 1740
DB 1772 TCAAGCTTCATCAGATGTTTGCAGCAGCAGGAAAGCGCAAGGTTGTGGCAG 1831
OY 1741 CTGATTAAGATCAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 1800
DB 1832 CTGATTAAGATCAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 1891
OY 1801 TTGATCTGAGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 1860
DB 1892 TTGATCTGAGAGAGATGAGAGATGAGAGATGAGAGATGAGAGATG 1951
OY 1861 CACATGAAGATGATGA 1878
DB 1952 CACATGAAGATGATGA 1969

RESULT 4
AF239719 2254 bp DNA linear PLN 02-JUN-2000
LOCUS Arabidopsis thaliana SGS3 gene, complete cds.
DEFINITION AF239719
ACCESSION AF239719
VERSION AF239719.1 GI:8164029
KEYWORDS
SOURCE Arabidopsis thaliana (chale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsie.
1 (bases 1 to 2254)
Mourrain, P., Beclin, C., Elmayan, T., Feuerbach, F., Godon, C.,
Morel, J.B., Jonette, D., Jacombe, A.M., Nikic, S., Picault, N.,
Remoue, K., Sanial, M., Vo, T.A. and Vaucheret, H.
Arabidopsis SGS2 and SGS3 genes are required for
posttranscriptional gene silencing and natural virus resistance
Cell 101 (5), 533-542 (2000)

TITLE
JOURNAL MEDLINE 20306668
PUBMED 10850495
2 (bases 1 to 2254)
Beclin, C., Mourrain, P., Vaucheret, H. and Elmayan, T.
Direct Submission
Submitted (28-FEB-2000) Biologie Cellulaire, INRA, Route de
Saint-Cyr, Versailles 78026, France

FEATURES
source
1. .2254
/organism="Arabidopsis thaliana"
/mol_type="genomic DNA"
/cultivar="Columbia"
/db_xref="taxon:3702"
/Chromosome="V"
/join(1..963,1037..1328,1440..1684,1787..1953,2044..2254)
/product="SGS3"
/join(1..963,1037..1328,1440..1684,1787..1953,2044..2254)
/codon_start=1
/product="SGS3"
/protein_id="AA073960.1"
/db_xref="GI:8164030"
/translation="MSSRAGPMKSKKVVQGYRPEVQLVQGLAGRLASODDGGEM
EVLISKRNKRGNTSGKTWVSQNSNPRAQGGQGGGSGVSRGNVSGRNGRG
IOANISGRGALSRKYDNNEVAPPPVPELEGWMAQAGSGAQTAVOEPDVED
VDNAEENDSDALDDSDDDSLADSDYSDVSQSHSRKQKWFKFPSLSLSLEQ
INERORMHCPAOCNGAIDWYLIHPLHARTGARRVYKARELAELVLEKDLQMG
ASVTPCEGRTYGVQWGLGDEKDYETVWPVVIINTPLDQNDKMLGKNGOELFE
DYELARRSHSYGQGRGMSVIMFESSATGYEAEHLHELMGLDRIAMQKRSK
FSGVRLQYGLPLATKDDLIFNHSQKTLKPLKLSYQGMVVKELRQISENQCLNY

```

```

ORIGIN
Query Match 77.8%; Score 1462; DB 8; Length 2254;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;
D*
FNNLSKONKGAUVLEESLEIMSSEKRLRTAEENRIVOROTMOHEONREMDAHDRE
MDSIQIHERDADKEENEMLOOQERAKVVOOQANINPSNDCCRRAEVSSFLFE
QKENEFEVEREERMLIKQEKEMDKKRHHIEIFDLEKEFDLEJOLMYKHGLHND
1 ATGAGTTCTAGGGCTGCTCCAAATGCTTAAGAGAAAGAAAGCTTCAAGGTGTTATAGGCT 60
1 ATGAGTTCTAGGGCTGCTCCAAATGCTTAAGAGAAAGAAAGCTTCAAGGTGTTATAGGCT 60
61 GAGTTGAACAGTTGGTTCAAGTTTGGCAGGACGAGCTGCTTTTCAAGATGAT 120
61 GAGTTGAACAGTTGGTTCAAGTTTGGCAGGACGAGCTGCTTTTCAAGATGAT 120
DB 61 GAGTTGAACAGTTGGTTCAAGTTTGGCAGGACGAGCTGCTTTTCAAGATGAT 120
121 GAGAGAGAGTGGAGAGTCAATTTCCAAAGAACACAGAACAAACCAAGAAACCTTCTGA 180
DB 121 GAGAGAGAGTGGAGAGTCAATTTCCAAAGAACACAGAACAAACCAAGAAACCTTCTGA 180
121 GAGAGAGAGTGGAGAGTCAATTTCCAAAGAACACAGAACAAACCAAGAAACCTTCTGA 180
DB 121 GAGAGAGAGTGGAGAGTCAATTTCCAAAGAACACAGAACAAACCAAGAAACCTTCTGA 180
181 AAAAATTGGGTTCTCAGATTTGGAATTCCTTAGAGCTTGGGGTGTGACAGCAAGG 240
DB 181 AAAAATTGGGTTCTCAGATTTGGAATTCCTTAGAGCTTGGGGTGTGACAGCAAGG 240
181 AAAAATTGGGTTCTCAGATTTGGAATTCCTTAGAGCTTGGGGTGTGACAGCAAGG 240
DB 181 AAAAATTGGGTTCTCAGATTTGGAATTCCTTAGAGCTTGGGGTGTGACAGCAAGG 240
241 AGAGGTAGCAACGATCTGGAGAGAGAACAAATGATCCGGAGAGGTAAACGGCAATGCT 300
DB 241 AGAGGTAGCAACGATCTGGAGAGAGAACAAATGATCCGGAGAGGTAAACGGCAATGCT 300
301 CGGGGCAATTCAGCTAACATATCTGTCGCGGAGAGAGCTTTGAGCAGAAAGTATGATAAC 360
DB 301 CGGGGCAATTCAGCTAACATATCTGTCGCGGAGAGAGCTTTGAGCAGAAAGTATGATAAC 360
301 CGGGGCAATTCAGCTAACATATCTGTCGCGGAGAGAGCTTTGAGCAGAAAGTATGATAAC 360
DB 301 CGGGGCAATTCAGCTAACATATCTGTCGCGGAGAGAGCTTTGAGCAGAAAGTATGATAAC 360
361 AACTTTGTGACACCCCACTGATCTCGCCCTCTTTGAGAGAGATGGAATTGGCAG 420
DB 361 AACTTTGTGACACCCCACTGATCTCGCCCTCTTTGAGAGAGATGGAATTGGCAG 420
361 AACTTTGTGACACCCCACTGATCTCGCCCTCTTTGAGAGAGATGGAATTGGCAG 420
DB 361 AACTTTGTGACACCCCACTGATCTCGCCCTCTTTGAGAGAGATGGAATTGGCAG 420
421 GCAAGAGAGAGTGTCTGCTCAGACACACAGCTGTGAGAGAGTTCCTGACGTGAGATGAT 480
DB 421 GCAAGAGAGAGTGTCTGCTCAGACACACAGCTGTGAGAGAGTTCCTGACGTGAGATGAT 480
421 GCAAGAGAGAGTGTCTGCTCAGACACACAGCTGTGAGAGAGTTCCTGACGTGAGATGAT 480
DB 421 GCAAGAGAGAGTGTCTGCTCAGACACACAGCTGTGAGAGAGTTCCTGACGTGAGATGAT 480
481 GGGATTAATGCTTCTGAGAGAGAGATGATCCGATGCTTTGATGATTCGATGACAC 540
DB 481 GGGATTAATGCTTCTGAGAGAGAGATGATCCGATGCTTTGATGATTCGATGACAC 540
481 GGGATTAATGCTTCTGAGAGAGAGATGATCCGATGCTTTGATGATTCGATGACAC 540
DB 481 GGGATTAATGCTTCTGAGAGAGAGATGATCCGATGCTTTGATGATTCGATGACAC 540
541 CTTGCAAGATGATTTATGATCTCGATGTGATGATGATGATGATGATGATGATGATGAT 600
DB 541 CTTGCAAGATGATTTATGATCTCGATGTGATGATGATGATGATGATGATGATGATGAT 600
541 CTTGCAAGATGATTTATGATCTCGATGTGATGATGATGATGATGATGATGATGATGAT 600
DB 541 CTTGCAAGATGATTTATGATCTCGATGTGATGATGATGATGATGATGATGATGATGAT 600
601 AATTAAGTGTCAAAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGATGATGAT 660
DB 601 AATTAAGTGTCAAAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGATGATGATGAT 660
601 AATTAAGTGTCAAAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGATGATGATGAT 660
DB 601 AATTAAGTGTCAAAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGATGATGATGAT 660
661 GAACACAGAGAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 GAACACAGAGAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
661 GAACACAGAGAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
DB 661 GAACACAGAGAGAGTGGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 720
721 TATTAAGTGTCAAAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGATGATGATGAT 780
DB 721 TATTAAGTGTCAAAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGATGATGATGAT 780
721 TATTAAGTGTCAAAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGATGATGATGAT 780
DB 721 TATTAAGTGTCAAAAAGTTCTTGGCAGCTTGAATGATGATGATGATGATGATGATGATGAT 780
781 CATAGAGATTTGGCTGAAGTTTGAAGAAAGATCTACAGATGAGAGAGCGCATCTGCATT 840
DB 781 CATAGAGATTTGGCTGAAGTTTGAAGAAAGATCTACAGATGAGAGAGCGCATCTGCATT 840
781 CATAGAGATTTGGCTGAAGTTTGAAGAAAGATCTACAGATGAGAGAGCGCATCTGCATT 840
DB 781 CATAGAGATTTGGCTGAAGTTTGAAGAAAGATCTACAGATGAGAGAGCGCATCTGCATT 840
841 CTTGTGTGATGATTTATGAGCAGATGAGAGAGTGGTGGTGGATGAGAAAGATTTATGA 900
DB 841 CTTGTGTGATGATTTATGAGCAGATGAGAGAGTGGTGGTGGATGAGAAAGATTTATGA 900
841 CTTGTGTGATGATTTATGAGCAGATGAGAGAGTGGTGGTGGATGAGAAAGATTTATGA 900
DB 841 CTTGTGTGATGATTTATGAGCAGATGAGAGAGTGGTGGTGGATGAGAAAGATTTATGA 900
901 ATTGTCTGGCTCCAAATGCTATCATCATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 ATTGTCTGGCTCCAAATGCTATCATCATGATGATGATGATGATGATGATGATGATGATGAT 960
901 ATTGTCTGGCTCCAAATGCTATCATCATGATGATGATGATGATGATGATGATGATGATGAT 960
DB 901 ATTGTCTGGCTCCAAATGCTATCATCATGATGATGATGATGATGATGATGATGATGATGAT 960

```

```

QY 961 A----- 961
Db 961 AAGTGAATCTCTCTCTTACTTCTTAATTTTCTCTGCAATTCATGATCTTA 1020
QY 962 -----AGTGGCTGGCATGGCCAAACCAAGAGCTGTGAAATCTTCGCAA 1007
Db 1021 GAATGTACATTTGATGGCTCGGCATGGGCCAAACCAAGAGCTGTGAAATCTTCGCAA 1080
QY 1008 GTATGAGGCTCTAGAGCAGCCATTCCTATGATCCACAGGCGCATGCTGGATGAGTGT 1067
Db 1081 GTATGAGGCTCTAGAGCAGCCATTCCTATGATCCACAGGCGCATGCTGGATGAGTGT 1140
QY 1068 TCTGATGTTGAGAGAGAGTCCCATGCTATTTGAGGCGCAAAGCTCCACCGGAGTT 1127
Db 1141 TCTGATGTTGAGAGAGAGTCCCATGCTATTTGAGGCGCAAAGCTCCACCGGAGTT 1200
QY 1128 AGCTGAGATGGGTTAGATGAATTTGCTGGGCTCAGAGCGCATATGTTTTCTGGAGG 1187
Db 1201 AGCTGAGATGGGTTAGATGAATTTGCTGGGCTCAGAGCGCATATGTTTTCTGGAGG 1260
QY 1188 TGTTCGCCAATGATAGGCTTCTTGCAGAGCAAGAGATCTGACATATTCATCAACA 1247
Db 1261 TGTTCGCCAATGATAGGCTTCTTGCAGAGCAAGAGATCTGACATATTCATCAACA 1320
QY 1248 CTCTCA----- 1253
Db 1321 CTCTCAAGTCTCTCCCAAAAGAAATTTGATATATGCTTTAGTTTTGTCATGGAAT 1380
QY 1254 -----AGG 1256
Db 1381 TTAAGTTTGTGTCCGTTATATGATCATCTGTTATATATATCATGATTCATTAGG 1440
QY 1257 CAAAACAGGCTGAATTTGAGTTGAATATATCCAAAGAGATGGTTGTAAGAGCTGAG 1316
Db 1441 CAAAACAGGCTGAATTTGAGTTGAATATATCCAAAGAGATGGTTGTAAGAGCTGAG 1500
QY 1317 GCAGATCTCTGAGAGCAATCAGCAGCTGAATCTCTTAAAGAAAGCTCAAAACGAA 1376
Db 1501 GCAGATCTCTGAGAGCAATCAGCAGCTGAATCTCTTAAAGAAAGCTCAAAACGAA 1560
QY 1377 CAAGCAGCCCAAGGCTCTGAGGAATCTCTGAAATTTATAGGAGAGAGCTGGTGAAC 1436
Db 1561 CAAGCAGCCCAAGGCTCTGAGGAATCTCTGAAATTTATAGGAGAGAGCTGGTGAAC 1620
QY 1437 TCGAGAGATATTCGATCTGTGAGACAGAGAACTAAGATCAGATGAAGAAACAGGA 1496
Db 1621 TCGAGAGATATTCGATCTGTGAGACAGAGAACTAAGATCAGATGAAGAAACAGGA 1680
QY 1497 AG----- 1498
Db 1681 AGAGGATGATTTTCTAGAAAATCACAAACTTGACATTTGTATTAAGTCTGATTC 1740
QY 1499 -----AGATGATGACACGA 1514
Db 1741 CATTTTGAATATATTTGCCAACAAAAAAGCTGTGGTGTGATGAGTGAATGATCACAACA 1800
QY 1515 CAGGTTTTTCATGATTTCAATCAACAGATCCATGAAAGAAAGAGCAAGAGAGAA 1574
Db 1801 CAGGTTTTTCATGATTTCAATCAACAGATCCATGAAAGAAAGAGCAAGAGAGAA 1860
QY 1575 TTTGAGATGTTGACAGAGAGAGAGTGCCAAGGTTGTTGGCAGAGACAGAGAAAT 1634
Db 1861 TTTGAGATGTTGACAGAGAGAGAGTGCCAAGGTTGTTGGCAGAGACAGAGAAAT 1920
QY 1635 TAATCCCTCTAGCAATGACGATTCGGAAG----- 1665
Db 1921 TAATCCCTCTAGCAATGACGATTCGGAAGAGATATATGTAATAATATCCCT 1980
QY 1666 ----- 1665
Db 1981 CTGGCGTTTTTTGTTTCAAACTTAAGAGTAATTAATCCGTTTTGATCTTTG 2040

```

```

QY 1666 -AGAGCTGAGAGAGTGTCAAGCTTTCATCGATTTCAAGAGAAAGAGATGAGAGTTTGT 1724
Db 2041 CAGAGCTGAGAGAGTGTCAAGCTTTCATCGATTTCAAGAGAAAGAGATGAGAGTTTGT 2100
QY 1725 GGAAGAGAGGAGATCTGTATTAAGATCAAGAGAAAGAGATGAGAGACATGAAGAGAG 1784
Db 2101 GGAAGAGAGGAGATCTGTATTAAGATCAAGAGAAAGAGATGAGAGACATGAAGAGAG 2160
QY 1785 GCATCAAGAGAGATTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAAACAGCTCAT 1844
Db 2161 GCATCAAGAGAGATTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAAACAGCTCAT 2220
QY 1845 GTACAGCATGGCTTCACATGAGATGATTTGA 1878
Db 2221 GTACAGCATGGCTTCACATGAGATGATTTGA 2254

RESULT 5
AX078760 3275 bp DNA linear PAT 22-FEB-2001
LOCUS
DEFINITION
Sequence 1 from Patent WO0105951.
ACCESSION
AX078760
VERSION
AX078760.1 GI:13158379
KEYWORDS
Arabidopsis thaliana (thale cress)
ORGANISM
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1

REFERENCE
1
Beclin,C., Elmayan,T. and Vaucheret,H.
Novel sgs3 plant gene and use thereof
Patent: WO 0105951-A 1.25-JAN-2001;
AVENTIS CROPS SCIENCE S.A. (FR) ; INSTITUT NATIONAL DE LA RECHERCHE
AGRONOMIQUE (FR)
FEATURES
Location/Qualifiers
source
1..3275
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"
primer_bind
693..715
/note="p356AD"
primer_bind
complement(2926)..2952
primer_bind
/note="p356Y"

ORIGIN
Query Match 77.8%; Score 1462; DB 6; Length 3275;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;

1 ATGAGTTCTTAGGGCTGCTCCATGCTTAAGAGAAAGAAAGCTTCAAGGTGTTATAGGCTT 60
Db 696 ATGAGTTCTTAGGGCTGCTCCATGCTTAAGAGAAAGAAAGCTTCAAGGTGTTATAGGCTT 755
QY 61 GAGGTGAACAGTGTGTTCAAGTTTGGCAGGAGCAGAGCTGGCTTTCACAAGATGAT 120
Db 756 GAGGTGAACAGTGTGTTCAAGTTTGGCAGGAGCAGAGCTGGCTTTCACAAGATGAT 815
QY 121 GAGAGAGAGTGGAGGTCTATTTCCAAAGAAACAAAGAACAAACAGAGAAACCTTTGGA 180
Db 816 GAGAGAGAGTGGAGGTCTATTTCCAAAGAAACAAAGAACAAACAGAGAAACCTTTGGA 875
QY 181 AAAAATTGGGTTTTCAGAAATTCGAATCTCTTAAGCTTTGGGGTGTGACAGCAAGGG 240
Db 876 AAAAATTGGGTTTTCAGAAATTCGAATCTCTTAAGCTTTGGGGTGTGACAGCAAGGG 935
QY 241 AGAGGTGAACAGTATCTGGAGAGAGAAACAATGTATCCGAGAGAGTAAACGCAATGCT 300
Db 936 AGAGGTGAACAGTATCTGGAGAGAGAAACAATGTATCCGAGAGAGTAAACGCAATGCT 995
QY 301 CGGGGCAATTCAGCTPAACATATCTGGTCGGGAGCAGAGCGTTGAGCAAGAAATGATTAAC 360
Db 996 CGGGGCAATTCAGCTPAACATATCTGGTCGGGAGCAGAGCGTTGAGCAAGAAATGATTAAC 1055

```

QY 361 AACTTGTGGCACCCCCACTGTATCTGCGCTCTCTTGGAGAGATGGAATTGGCAG 420
 Db 1056 AACTTGTGGCACCCCCACTGTATCTGCGCTCTCTTGGAGAGATGGAATTGGCAG 1115
 QY 421 GCAGAGAGAGGTTTCTGCTGAGACACAGCTGTGCGAGGATTTCTGAGCGTGAAGATAT 480
 Db 1116 GCAGAGAGAGGTTTCTGCTGAGACACAGCTGTGCGAGGATTTCTGAGCGTGAAGATAT 1175
 QY 481 GTGATTAATGCTTCTGAGAGAGATGAATTCGATGCTTTGATGATTTCTGATGACGAC 540
 Db 1176 GTGATTAATGCTTCTGAGAGAGATGAATTCGATGCTTTGATGATTTCTGATGACGAC 1235
 QY 541 CTTCGAGATGATGATTAATGATCTGCGATGTGAGTCAAAAAGCCATGATCAAGAAACGAG 600
 Db 1236 CTTCGAGATGATGATTAATGATCTGCGATGTGAGTCAAAAAGCCATGATCAAGAAACGAG 1295
 QY 601 AATTAAGTGTGCAAAAAGTCTTTGGCAGCTTGATGATGCTTGATGATGATGATGATGAT 660
 Db 1296 AATTAAGTGTGCAAAAAGTCTTTGGCAGCTTGATGATGCTTGATGATGATGATGATGAT 1355
 QY 661 GAACCAAGAGGAGGATGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 720
 Db 1356 GAACCAAGAGGAGGATGAGATGCTGATGATGATGATGATGATGATGATGATGATGATGAT 1415
 QY 721 TATTAACCTGCACTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
 Db 1416 TATTAACCTGCACTCTCTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1475
 QY 781 CATAGAATTTGGCTGAAGTTTGAAGAAAGATCTACAGATGAGAGCGCATGTCAT 840
 Db 1476 CATAGAATTTGGCTGAAGTTTGAAGAAAGATCTACAGATGAGAGCGCATGTCAT 1535
 QY 841 CCTTGTGTGAGATTATAGGAGAGTGAAGGGTTTGGGTGAGATGAAAAGATTATGAA 900
 Db 1536 CCTTGTGTGAGATTATAGGAGAGTGAAGGGTTTGGGTGAGATGAAAAGATTATGAA 1595
 QY 901 ATTGTCTGGCTCCATATGATCATCATATGATTAATTAAGATTAAGAGCGATTAACGAT 960
 Db 1596 ATTGTCTGGCTCCATATGATCATCATATGATTAATTAAGATTAAGAGCGATTAACGAT 1655
 QY 961 A----- 961
 Db 1656 AAGGTGAATTTCTTCTCTTTAATTTTCTCTTCAATTTACTGATCTTA 1715
 QY 962 -----AGTGGCTCGCATGCGAACCAAGAGCTGTGGAATTTCTTCAACA 1007
 Db 1716 GAATGTTACATTGATGAGCTCGGCATGAGGCAACCAAGAGCTGTGGAATTTCTTCAACA 1775
 QY 1008 GTATGAGGCTCTTAAGACAGGCGCATTTCTATGATGCAAGGCGCATGCTGAGATGAGTGT 1067
 Db 1776 GTATGAGGCTCTTAAGACAGGCGCATTTCTATGATGCAAGGCGCATGCTGAGATGAGTGT 1835
 QY 1068 TCTGATGTTTGAAGAGAGTGCCTATTTTGAAGCGCGAACGCTTCAACGGAGGTT 1127
 Db 1836 TCTGATGTTTGAAGAGAGTGCCTATTTTGAAGCGCGAACGCTTCAACGGAGGTT 1895
 QY 1128 AGCTGAAGTGGGTTAATTAAGATTTGCTGGGGTCAAGAGCGCATGATTTTCTGAGG 1187
 Db 1896 AGCTGAAGTGGGTTAATTAAGATTTGCTGGGGTCAAGAGCGCATGATTTTCTGAGG 1955
 QY 1188 TGTTCGCAACTGATGAGCTTCTTGCACAGACAGATCTGAGCATATTTCAATCAACA 1247
 Db 1956 TGTTCGCAACTGATGAGCTTCTTGCACAGACAGATCTGAGCATATTTCAATCAACA 2015
 QY 1248 CTCTCA----- 1253
 Db 2016 CTCTCAAGGTTCTCTCCCAAGAAATTTGATATATGCTTTAGTTTGTCAATGGAAT 2075
 QY 1254 -----AGG 1256
 Db 2076 TTAAGATTGTGTGCTGCTGATATGATCTGTTAGTATATATCTATGATTCATTTAGG 2135

QY 1257 CAAAACAGGCTGAATTTGAGTTGAATATATCAAGAGATGTTGTAAAGAGCTGAG 1316
 Db 2136 CAAAACAGGCTGAATTTGAGTTGAATATATCAAGAGATGTTGTAAAGAGCTGAG 2195
 QY 1317 GCAGATCTCTGAGGACATTCAGAGCTGAACTTATTAAGAACAGCTCTCAAAACGAA 1376
 Db 2196 GCAGATCTCTGAGGACATTCAGAGCTGAACTTATTAAGAACAGCTCTCAAAACGAA 2255
 QY 1377 CAACGACCCCAAGGCTTGAAGAAATCTGGAATTTATGAGCGAGAGAGCTGTAGAAC 1436
 Db 2256 CAACGACCCCAAGGCTTGAAGAAATCTGGAATTTATGAGCGAGAGAGCTGTAGAAC 2315
 QY 1437 TGCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1496
 Db 2316 TGCAGAGATTAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2375
 QY 1497 AG----- 1498
 Db 2376 AGAGGTATGATTTTCTTGAAGAAATCAACAACTTGACATTTTGTATTACTTACTGATTTCA 2435
 QY 1499 -----AGATGATGACACGA 1514
 Db 2436 CATTTTGAATTAATTTCTTCAACAAACAACTGATGATTTTGAAGATGATGACACGA 2495
 QY 1515 CAGTTTTCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 1574
 Db 2496 CAGTTTTCATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 2555
 QY 1575 TTTGAGATGTTGCGAGCGAGGAACTGCGCAAGTTTGGCCAGCGAGCGAGGAACT 1634
 Db 2556 TTTGAGATGTTGCGAGCGAGGAACTGCGCAAGTTTGGCCAGCGAGCGAGGAACT 2615
 QY 1635 TAAATCCCTAGCATGACGATGCGGAAAG----- 1665
 Db 2616 TAAATCCCTAGCATGACGATGCGGAAAGAGATATGTAATTAATTAATTAATTAATTAAT 2675
 QY 1666 ----- 1665
 Db 2676 CTGGGCTTTTGTTCCTTCAACCTAAGATTAATTAATTAATTAATTAATTAATTAATTAAT 2735
 QY 1666 -AGAGTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1724
 Db 2736 CAGAGCTGAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2795
 QY 1725 GGAAGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1784
 Db 2796 GGAAGAGAGGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2855
 QY 1785 GCATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1844
 Db 2856 GCATCAAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 2915
 QY 1845 GTTAAAGCATGAGCTTCAATGAAGATGATGGA 1878
 Db 2916 GTTAAAGCATGAGCTTCAATGAAGATGATGGA 2949

RESULT 6
 AB025633 81365 bp DNA linear PIN 27-DEC-2000
 LOCUS Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MGM1.
 DEFINITION Arabidopsis thaliana genomic DNA, chromosome 5, P1 clone:MGM1.
 ACCESSION AB025633 Ba000015
 VERSION AB025633.2 GI:10178221
 KEYWORDS
 SOURCE
 ORGANISM Arabidopsis thaliana (thale cress)
 Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
 1 (bases)
 Sato,S., Nakamura,Y., Kaneko,T., Katoh,T., Asamizu,E., Kocani,H.
 and Tabata,S.
 Structural analysis of Arabidopsis thaliana chromosome 5. X.


```

/codon_start=1
/evidence=not experimental
/protein_id="BAA97244.1"
/db_xref="GI:8809703"
/translation="MSSRAGPMSKERKQVGYRPEVEEDLVGLACTRLASQDDGGEG
EYISKKNKPKNTGKTIVSQNSNPSPRMWGQOQGRSNVSGRNNVSGKNGRG
IOANTSGRRLSKRYKDNFVAPPVSPLEGGMWQARQSGAQTHTPEFVDEED
VUNAISREEDSALDDDDDDLDYDSDVSQSHSGKQNMKEKFPFGLDLSIEQ
INERORHICPAQCNQGRGADIMWYNLPLHARTKGRVYKRLHRELALEKDLQMRG
ASVTPGEBYGGKIGEDKDEKDEYELVWPMVYIMNTPLDKDNDKMLGMOELLEYF
DKKELARHSHYGPQGRHMSVLMESSTGYLEBRHLRELAEMGLDIRIAMGKRSV
FSGGVOLYGLATKQDDLEFNQSHQKTRLFELKSIOENVVKELRQISENQDLY
FKRLSKONKRAKYLEESLEIMSEKRLRTAEDNRIYRQTKWQONRBNEMADHFRF
MSIKOIHERRDAKEENFEMLOQOERAKVGOQOONINSSNDRCRKAEEVSSIEF
QEKEMEFYERHEMLIKDEKQEDMKHHEIEIPDEKEFDEALEQLMYKGLHNEED
D"
join(25614..25838,26061..26207,26324..26438,26516..26610,
26710..26813,27522..27765,27943..28036,28115..28305,
28385..28604,28688..28855,28944..29092)
/notes="gb|AAC97420.1
gene_id:MOM1.16
strong similarity to unknown protein"
/codon_start=1
/evidence=not experimental
/protein_id="BAA97243.1"
/db_xref="GI:8809702"
/translation="MAPPAQQTAAVAEAAGADAPPOQGRFGSTISGIVRIAVWFY
ASKPSPKQMDPTAPSOAMTNI.FHKGESLDWMFYLSQEKENDGNDRALYHETN
IPVAVTPBSIRTKSLIYVPSSETLONNGSLYAHIPARSGFPIDPEVDEPQPLNSFR
THAVATFPOKKNKKSLLGSPKSDSESPVEKYGDKSDKEEVPWEMISLMEPN
VTINLVDFPQADLVEPTTGNTGYPTTYNREFMLRDKRIPNENVESELPLNLEISFV
ISWAKMQLFOQVQDSFQWQSYSGMSLDESDELKFLPBLDITMFTMSLSHVF
DPLAFKNDIOFMNKNKMEBLSASVVLNICOVFLFLYLDNDTSMILASGVVC
IEFWKIKAMRIEVDKSGMIPLRFLHREBSYNSKKEYDIAIKELSYVLLLVGL
SYSLAYERRHKSWSMILSISLTCVYFIMMCPOLFIYKLSVAHLPMWMTYGL
LNTIIDLFAFVIMPLIHLRHSVRDVIPLIYLQWVYPVDKTRVNEGFGEDEET
AEKRLTEKEBEDKTN"
complement(join(29439..29576,29661..29885,29985..30268,
30341..30493,30579..30722,30956..31484))
/notes="gene_id:MOM1.15
77.8%; Score 1462; DB 8; Length 81365;
Best Local Similarity 83.3%; Pred. No. 0;
Matches 1878; Conservative 0; Mismatches 0; Indels 376; Gaps 4;

```

```

QY 421 GCAGAGAGAGGTTCTGCTACAGACACAGCGTGTGACAGATTCTCTGACGTGAGATGAT 480
DB 23092 GCAAGAGAGAGGTTCTGCTACAGACACAGCTGTGACAGATTCTCTGACGTGAGATGAT 23151
QY 481 GTGGATTAATGCTTCTGAGAGAGAAATGATTCGATGCTTTGGATGATTTCTGATGACGAC 540
DB 23152 GTGATTAATGCTTCTGAGAGAGAAATGATTCGATGCTTTGGATGATTTCTGATGACGAC 23211
QY 541 CTTCGAAGTATGATTTATGATCTGGAGTGTGAGTCGAAAGAGCCATGATATCAAGAAAGCAG 600
DB 23212 CTTCGAAGTATGATTTATGATCTGGAGTGTGAGTCGAAAGAGCCATGATATCAAGAAAGCAG 23271
QY 601 AATAAGTGTTCAAAAGAGTCTTTGGCAGCTTGATAGCTTTGATCGAGCAGATTAAT 660
DB 23272 AATAAGTGTTCAAAAGAGTCTTTGGCAGCTTGATAGCTTTGATCGAGCAGATTAAT 23331
QY 661 GAACCAAGAGGAGTGGCATTTGTCAGCTTTGTCAGAACGAGCACTGGTGCATGCAATGG 720
DB 23332 GAACCAAGAGGAGTGGCATTTGTCAGCTTTGTCAGAACGAGCACTGGTGCATGCAATGG 23391
QY 721 TATAACCTGCACCTCTACTAGCTCATGCGAGCAAAAGAGCTAGCGAGTTAAGCTC 780
DB 23392 TATAACCTGCACCTCTACTAGCTCATGCGAGCAAAAGAGCTAGCGAGTTAAGCTC 23451
QY 781 CATAGAGAAATGGCTGAAGTTTAAAGAAAGATCTACAGATGAGAGCCGATCTGTCAAT 840
DB 23452 CATAGAGAAATGGCTGAAGTTTAAAGAAAGATCTACAGATGAGAGCCGATCTGTCAAT 23511
QY 841 CCTGTGTGAGATTTATAGGCACTGGAAGGTTTGGGTGAGATGAAAGAGATTAAGAA 900
DB 23512 CCTGTGTGAGATTTATAGGCACTGGAAGGTTTGGGTGAGATGAAAGAGATTAAGAA 23571
QY 901 ATGTCTGAGCTTCAATGATGATCATCATGATATCTAGACTGATTAAGAGCAGATACGAT 960
DB 23572 ATGTCTGAGCTTCAATGATGATCATCATGATATCTAGACTGATTAAGAGCAGATACGAT 23631
QY 961 A----- 961
DB 23632 AAGTGAATCTCTGCTCTTTTACTTTTAAATTTTCTCTTGCAATCTGATCTTA 23691
QY 962 -----AGTGGCTGGCATGGGCAACCAAGAGCTGTGAATATCTTGACAA 1007
DB 23692 GAATGTACATTTGATGCTGTGCTGCGATGGCAACCAAGAGCTGTGAATATCTTGACAA 23751
QY 1008 GTATGAGGCTCTTAAGACACGCAATTCCTATGCTCAAGGCGCATCTGGAGATGAGGT 1067
DB 23752 GTATGAGGCTCTTAAGACACGCAATTCCTATGCTCAAGGCGCATCTGGAGATGAGGT 23811
QY 1068 TCTGATGTTGAGAGCAGTGCATCTGGCTATTGGAGGCCGAAAGCTCTCACCGGAGTT 1127
DB 23812 TCTGATGTTGAGAGCAGTGCATCTGGCTATTGGAGGCCGAAAGCTCTCACCGGAGTT 23871
QY 1128 AGCTGAGAGGGGTTAGATAGATTCCTGGGGTCAAGAACGCAATGTTTCTTGAGG 1187
DB 23872 AGCTGAGAGGGGTTAGATAGATTCCTGGGGTCAAGAACGCAATGTTTCTTGAGG 23931
QY 1188 TGTTCGCAACTGATAGCTTCTTTCACAGAGCAAGATCTGGACATTTCAATCAACA 1247
DB 23932 TGTTCGCAACTGATAGCTTCTTTCACAGAGCAAGATCTGGACATTTCAATCAACA 23991
QY 1248 CTCTCA----- 1253
DB 23992 CTCTCAAGGTTCTCTCCCAAGAAATTTGATATATGCTTTAGTTTGTCAATGAAT 24051
QY 1254 -----AGG 1256
DB 24052 TTAAGTTTGTGTCGCGTTAATGATCTGTATATATATATATATATATATATATATGAT 24111
QY 1257 CAATAAGAGCTGAATTTGAGTTGAATATATCAAGAGATGTTGTAAGAGCTGAG 1316
DB 24112 CAATAAGAGCTGAATTTGAGTTGAATATATCAAGAGATGTTGTAAGAGCTGAG 24171
QY 1317 GCAGATCTGAGAGCAATCAGAGCTGAATCTACTTTAAGAACAGCTCTCAAAAAGAA 1376

```

```

Db      24172 GCAGATCTGAGAGACATCAGACGCTGAACTCTTAAAGAACAGCTCAAAACAGAA 24231
QY      1377  CAGACGCGCAAGGTGCTGAGAAATCTGGAATTAATAGAGAGAGGTGGTGAAC 1436
Db      24232 CAGACGCGCAAGGTGCTGAGAAATCTGGAATTAATAGAGAGAGGTGGTGAAC 24291
QY      1437  TGCAGAGATATTCGATCGTGAAGACAGAACTTAAGATCAGCATGAACAGAGGGA 1496
Db      24292 TGCAGAGATATTCGATCGTGAAGACAGAACTTAAGATCAGCATGAACAGAGGGA 24351
QY      1497  AG----- 1498
Db      24352 AGAGATGATATTTCTAGAAATATCAAACTTACATTTGATTAATTAATTAATTA 24411
QY      1499  ----- 1499
Db      24412 CATTTTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 24471
QY      1515  CAGGTTTTCATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1574
Db      24472 CAGGTTTTCATGATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 24531
QY      1575  TTTGAGATGTTTCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1634
Db      24532 TTTGAGATGTTTCAAGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24591
QY      1635  TAATCCCTCTAGCAATGACATGCTCCGAAAG----- 1665
Db      24592 TAATCCCTCTAGCAATGACATGCTCCGAAAGAGATTAATTAATTAATTAATTAATTA 24651
QY      1666  ----- 1665
Db      24652 CTGGCGTTTGTGTTTTCCTCAAACTTAAGATTAATTAATTAATTAATTAATTAATTAATTA 24711
QY      1666  AGAGCTGAGAAAGTGTCAAGCTTCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 1724
Db      24712 CAGAGCTGAGAAAGTGTCAAGCTTCAATGATTAATTAATTAATTAATTAATTAATTAATTAATTA 24771
QY      1725  GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1784
Db      24772 GGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24831
QY      1785  GCATCAAGAGAGAGATTTGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1844
Db      24832 GCATCAAGAGAGAGATTTGATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 24891
QY      1845  GTACAAAGATGAGCTTCAAGATGAAGATGA 1878
Db      24892 GTACAAAGATGAGCTTCAAGATGAAGATGA 24925

```

```

RESULT 7
ATHS28171 650 bp DNA linear PLN 29-MAR-2003
LOCUS      Arabidopsis thaliana T-DNA flanking sequence, left border, clone
DEFINITION
VERSION    AJ528171.1 GI:26796431
KEYWORDS   left border; T-DNA flanking sequence.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicotyledons;
            rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsids.
REFERENCE
AUTHORS    Brunaud, V., Balzerque, S., Dubreucq, B., Aubourg, S., Samson, F.,
            Chauvin, S., Bechtold, N., Criau, C., Depose, R., Pelletier, G.,
            Lepoint, L., Caboche, M. and Lecharny, A.
TITLE      T-DNA integration into the Arabidopsis genome depends on sequences
            of pre-insertion sites
JOURNAL    EMO Rep. 3 (12), 1152-1157 (2002)
MEDLINE    22363535

```

```

PUBMED 12446565
REFERENCE 2 (bases 1 to 650)
AUTHORS  Balzerque, S.
TITLE    Direct Submission
JOURNAL  Submitted (21-NOV-2002) Balzerque S., UMRGV, INRA/CNRS, 2 rue
            Gaston Cremieux, 91057 Evry cedex, FRANCE
COMMENT  PCR was performed on DNA from transformants of Arabidopsis thaliana
            plants from INRA (Versailles). The DNA fragment(s) resulting from
            the PCR were directly sequenced from the left or the right border
            to determine the genomic sequence flanking the insertion. T-DNA
            derived sequences were removed. Information to order the
            corresponding mutant line and a link to a database providing a
            graphical display of the insertion site are available at
            http://dbsgap.versailles.inra.fr/publiclines/. This sequence has
            been generated in the framework of the French plant genomics
            program 'Genoplante' (http://www.genoplante.com and
            http://genoplante-info.inbio.gen.fr).
            Location/Qualifiers
                source
                    1..650
                        /organism="Arabidopsis thaliana"
                        /mol_type="genomic DNA"
                        /culтивat="massillewskija"
                        /db_xref="taxon:3702"
                        /clone="157C08"
                        /clone_lib="Arabidopsis thaliana T-DNA insertion lines"
                misc_feature
                    1..650
                        /note="T-DNA flanking sequence
                        left border"
ORIGIN
Query Match 23.3%; Score 437.8; DB 8; Length 650;
Best Local Similarity 87.4%; Pred. No. 9e-96;
Matches 522; Conservative 0; Mismatches 2; Indels 73; Gaps 1;
QY 733 CCTTACTAGCTCATCGAGACAAAGAGAGCTAGGCGAGTTAAGTCCATAGAGAAATTG 792
Db 1 CCTTACTAGCTCATCGAGACAAAGAGAGCTAGGCGAGTTAAGTCCATAGAGAAATTG 60
QY 793 GCTGAGTTTGAAGAAAGATCTACAGATGAGAGGCGCATCTGCTTCTGGTGGAG 852
Db 61 GCTGAGTTTGAAGAAAGATCTACAGATGAGAGGCGCATCTGCTTCTGGTGGAG 120
QY 853 ATTATGAGCAGTGAAGGTTGGTGGTGAAGATGAAGAAAGATTAATTAATTAATTAATTAATTA 912
Db 121 ATTATGAGCAGTGAAGGTTGGTGGTGAAGATGAAGAAAGATTAATTAATTAATTAATTAATTA 180
QY 913 CCAATGTCATCATCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 961
Db 181 CCAATGTCATCATCTGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 240
QY 962 ----- 961
Db 241 TTCTGCTTTTAATCTTTAATTTTCTCTTGCAATTAATTAATTAATTAATTAATTAATTAATTA 300
QY 962 --AGTGGCTCGGAGTGGCAACCAAGAGCTGTGAATTAATTAATTAATTAATTAATTAATTAATTA 1019
Db 301 GTATGAGCTCGGAGTGGCAACCAAGAGCTGTGAATTAATTAATTAATTAATTAATTAATTAATTA 360
QY 1020 TAGAGACGCCATCTCTATGATTCACAGGAGCCATCTGTGAGATGATGTTCTGATGTTTGA 1079
Db 361 TAGAGACGCCATCTCTATGATTCACAGGAGCCATCTGTGAGATGATGTTCTGATGTTTGA 420
QY 1080 GAGCAATGCACTGGGTATTTGAGAGCCGAAAGCCCTCCACCGGAGATTAATTAATTAATTAATTA 1139
Db 421 GAGCAATGCACTGGGTATTTGAGAGCCGAAAGCCCTCCACCGGAGATTAATTAATTAATTAATTA 480
QY 1140 GTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 1199
Db 481 GTTATGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 540
QY 1200 GTATGAGCTTCTTGAACAAGAGATTTGAACATTAATTAATTAATTAATTAATTAATTAATTA 1256
Db 541 GTATGAGCTTCTTGAACAAGAGATTTGAACATTAATTAATTAATTAATTAATTAATTAATTA 597

```


QY 1057 GGGATGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1116
 |||||
 Db 1070 GGGATGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1129
 |||||
 QY 1117 CACCGGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1169
 |||||
 Db 1130 CAGATGATGTTTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1189
 |||||
 QY 1170 CAGATGATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1229
 |||||
 Db 1190 TACATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1249
 |||||
 QY 1230 GAGATGATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1289
 |||||
 Db 1250 GAGATGATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1309
 |||||
 QY 1290 CCAAGATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1349
 |||||
 Db 1310 TATGATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1369
 |||||
 QY 1350 CTTTGAAGATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1409
 |||||
 Db 1370 CTTTGAAGATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1429
 |||||
 QY 1410 AATTGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1469
 |||||
 Db 1430 TGTGTTTGTTCGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1489
 |||||
 QY 1470 TAAATGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1529
 |||||
 Db 1490 CAAAGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1549
 |||||
 QY 1530 TTAATGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1588
 |||||
 Db 1550 TCAATGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1609
 |||||
 QY 1589 AGAGAGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1648
 |||||
 Db 1610 TGAAGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1660
 |||||
 QY 1649 ATGAGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1708
 |||||
 Db 1661 CTGAAGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1720
 |||||
 QY 1709 AGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1768
 |||||
 Db 1721 ATGTTGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1780
 |||||
 QY 1769 AAGATGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1828
 |||||
 Db 1781 TGAAGTCAAGAGAGTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1840
 |||||
 QY 1829 CTTTGAAGATGTTTGTTCGAGAGTGTTCGATGTTTGGAGAGAGTGCCTGCTATTGAGAGCCGACGCTC 1853
 |||||
 Db 1841 CCCTCACTTCCCTGATGATTAACA 1865

RESULT 9
 AKI00699 2570 bp mRNA linear PLN 24-JUL-2003
 LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:J023114M02, full
 DEFINITION
 Accession AKI00699
 VERSION AKI00699.1 GI:32985908
 KEYWORDS
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Eriatoidae; Oryzaceae; Oryza.
 REFERENCE
 AUTHORS The Rice Full-length cDNA Consortium, National Institute of

TITLE

JOURNAL

MEDLINE

PUBMED

AUTHORS

2 (bases 1 to 2570)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K.,

Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W.,

Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T.,

Horii, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imanura, K.,

Imotani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I.,

Kanagawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M.,

Kikuchi, S., Kishikawa-Hirano, T., Kishimoto, N., Kobayashi, M.,

Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,

Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M.,

Masuda, H., Matsubara, K., Matsuyama, T., Mura, J., Miyazaki, A.,

Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M.,

Namiki, T., Narikawa, R., Nishikawa, R., Nishikawa, K., Oka, H.,

Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H.,

Osato, N., Ota, Y., Otomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K.,

Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K.,

Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S.,

Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y.,

Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A.,

Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W.,

Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S., and

Yoshimura, A.

Direct Submission

Submitted (27-Aug-2002) Shoshi Kikuchi, National Institute of

Agronomical Sciences, Department of Molecular Genetics, Head of

Laboratory of Gene Expression, 2-1-2 Kannondai, Tsukuba, Ibaraki

305-8602, Japan (E-mail: skikuchi@ias.affrc.go.jp,

Tel:81-29-838-7007, Fax:81-29-838-7007)

This clone is one of the 28k full-length cDNA clones from japonica

rice.

URL: <http://cdna01.dna.affrc.go.jp/cDNA/>

NIAS Rice Full-length cDNA Project Team: Kikuchi, S., Satoh, K.,

Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J.,

Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T.,

Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T., and

Yamamoto, M.

FALS Genome Sequencing & Analysis Group: Otomo, Y., Iida, Y.,

Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M.,

Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Mura, J.,

Mizuno, K., Narikawa, R., Nishikawa, R., Oka, M., Ryu, R., Sugano, S.,

Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S.,

Yoshimura, A., Matsubara, K., and Murakami, K.

Genome Exploration Research Group in Riken Genomic Sciences Center

and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K.,

Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T.,

Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K.,

Hirooka, T., Horii, F., Iida, J., Imanura, K., Imotani, K., Ishii, Y.,

Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J.,

Kishikawa-Hirano, T., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,

Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,

Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Osato, N.,

Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H.,

Sasakawa, T., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T.,

Shishiki, T., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M.,

COMMENT

Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shitaka, I.,
Sogade, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F.,
Takahara-Khanira, S., Tanaka, F., Tomaru, A., Toyota, T., Waki, K.,
Yasunishi, A. and Hayashizaki, Y.

Location/Qualifiers

1. 25/01/2025

/organism="Oryza sativa (japonica cultivar-group)"

/mol_tyr

/cultivar="Nippombare"

```
/db_xref="EAXon:3994/"
```

```
/clone="J0Z3114M0Z
```

ORIGIN

Query Match	19.8%;	Score 372;	DB 8;	Length 2300;
Best Local Similarity	57.0%;	Pred. No. 1e-79;		
Matches 786;	Conservative	0;	Mismatches 575;	Indels 19;
				Gaps 5

[illegible]

QY	1314	GAGCGAGATCTCTGAGGACAAATCAGAGGTGAACCTTATTAAGCAACAGCTCTCAAAACA	1373
Db	1546	GAGCGAGATAGTGAAGACAAACCAACTGATTTATGTGMAAGATGATGGTTAAGAC	1605
QY	1374	GAAACAAGCAGCCAGAGTGCTTGAGGAATCTCTGAAATTATGAGCGAAGACTGTGTG	1433
Db	1606	AGAGCCGACATTCCTAAAGATAGTAAAGAGCTTGAGGTGTGAAAGCCAGAAATTTAGGG	1665
QY	1434	AATCTCGAGAGATTAATCCGATCGTGAGACAGAGAACTAAGATCAGCATGAAACGAACG	1493
Db	1666	AGCAATTGAAGATATATTTATTTTGAAGAGAAAACCAAGAGATGCTCTCAGAGTGGCT	1725
QY	1494	GGAAGAGATGAGTGCACACGACGAGTTTTTCATGATTCATTAACAGATCCATGAAAG	1553
Db	1726	GGAACAGATGGAGTTTATCGAAATTTTATCATGAACAAATTTGAAGACTTTCGAAAGG	1785
QY	1554	AAGAGCGCAAGGAGAGGAAATTTCGAGATGTTGACGACAGAGAAAGTCCAAAGTTGT	1613
Db	1786	CACGAGAGAAAGAAATGATGTTGAGAGGTTCTGCAGAGAGGCTACAGAAAGCTAT	1845
QY	1614	TGGCCAGACGACGAGAACATTAAATCCCTCTAGCAATGACGATTCGCCGAAAGAGCTGA	1673
Db	1846	AG-----AGTGTGATGTTGACTCTGAAACTACGAAATTTGCAAGCTTAAGGAGGA	1896
QY	1674	GGAAGTGTCAAGCTTCATCGAGTTTCAAGAAAGAGATGAGAGAGTTTGTGAAAGAG	1733
Db	1897	ACAGTACAAAGATCATAGATTCGACAGGTGAAGATGCGGAGAGAGTTGATGTGTGAACA	1956
QY	1734	GGAGATGCTGATAAAGATCAAAGAGAAAGATGGAAAGCATGAAGAAAGAGGACTCAACA	1793
Db	1957	AGATTAATTTGATCAAAACACAGGAAGAAAGGCGAACTCAAGATGAGAGTATATGGC	2016
QY	1794	GGAGATATTTGATCTGAGAGAAAGATTTGATGAGGCTTTGGAACAGCTCATGTACAGCA	1853
Db	2017	AAAGATGTTGAGCTCGAGAGAGACTTTATGCCGCTCTCAACAGTCTGATGAGAGAAACA	2076

RESULT 10	AK064217	LOCUS	DEFINITION
AK064217	1556 bp	mRNA	linear
			PLN 24-JUL-2003
			CDNA clone:002-104-F07, full
			<i>Oryza sativa</i> (japonica cultivar-group)

accession AK064217.1 GI:32974735
 version F11_CDNA
 keywords Oryza sativa (japonica cultivar-group)
 source Oryza sativa (japonica cultivar-group)
 organism Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzaceae; Oryza.
 reference 1. Oryza sativa F11 research CDNA Consortium. National Institute of

REFERENCE	AUTHORS	TITLE
1	The Rice Full-length cDNA Consortium, National Institute of Agricultural Sciences Rice Full-length cDNA Project Team, Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Onitsuki, K., Shishikiri, T., Foundation of Advancement of International Science Genome Sequencing & Analysis Group: Ootomo, Y., Murakami, K., Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y., Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, O., Lu, M., Naitakawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Mitkura, J., Ikeda, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J., Kunumegi, T., Oka, M., Ryu, R., Ueda, M., Matubara, K., RIKEN, Kawai, J., Carrilho, P., Adachi, J., Aizawa, K., Arakawa, T., Fukuda, S., Kagawa, I., Kondo, S., Komoto, H., Miyazaki, A., Oono, N., Ota, Y., Saito, R., Sasaki, D., Seto, K., Shibata, K., Shinagawa, A., Shiraiki, T., Yoshino, M., and Hayashizaki, Y.	Japanese rice collection, mapping, and annotation of over 28,000 cDNA clones from
JOURNAL MEDLINE	22752273	
PUBMED	12869764	

RESULT 11
AF542974 1564 bp mRNA linear PLN 02-JUL-2003
LOCUS Triticum aestivum Emr1 mRNA, complete cds.
DEFINITION AF542974
VERSION AF542974.1 GI:32401385
KEYWORDS
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticaceae; Triticum.
AUTHORS Zhao, X., Li, Q. and Zhang, X.
TITLE Isolation and expression of a new kind of gene involve in
embryogenesis in Triticum aestivum L.
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1564)
AUTHORS Zhao, X., Li, Q. and Zhang, X.
TITLE Direct Submission
JOURNAL Submitted (03-SEP-2002) College of Life Sciences, Shandong
Agricultural University, Daizong Street, Tai'an, Shandong 271018,
P. R. China
FEATURES
Source
1. 1564
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultiivar="PH 82-2-2"
/db_xref="taxon:4565"
/dev_stage="embryo"
309..1223
/note="Involved in embryogenesis"
/codon_start=1
/product="Emr1"
/protein_id="AAP80862.1"
/db_xref="GI:32401386"
/translation="MGQELLEYSYDAATKARAYPGGHRGMSYLIFSSAVGWE
AERLKHFIIDORTDRTDNRRVPLPGKQLYGLFARKEDETHFRHOCGRPKY
EMRSHNEVAVKQKMSQEDNQNLVLRKVVTKRSRKVEETLGVITQKREMBN
IIVRSKAKKVEEEMKSOEKFPHQIENIHATDKESEFRLIQEERARQCD
VDSGTEERRRKRQVORFICQVDVQEFPAERDEMIKAIHEKRVQIKKEYAKVE
LEKFDALVTGLMEKRRPTFOASSSP"
CDS
Query Match 15.5%; Score 291.6; DB 8; Length 1564;
Best Local Similarity 59.2%; Pred. No. 4.2e-60;
Matches 539; Conservative 0; Mismatches 359; Indels 12; Gaps 2;
ORIGIN
952 GATTAACGATTAAGTGGCTGGCATGGGCAACCAAGAGCTGGGAATCTTGACAAAGTAT 1011
288 GATGCTCAGTAAAGTGGTAAAGGCAACCAAGAACTTCTGAGTATTCAGTCAATTAT 347
1012 GAGGCTCTTAGAGCAGCCGCTCTTAGTTCACAGGGCCATCTGGAGTGAAGTCTTG 1071
348 GCTGCGACCAAGACAGCCATGATATGTCAGAGTGGGACCGTGGCATGAGTGTGTA 407
1072 ATGTTTGAAGCAGTGGCCTGCTATTTGAGGCGGAAACGCTCCACCGGAGTAGCT 1131
408 ATATTTAAAGCTCTGCTGGGCTATATGAGGCAAGACGTTCTATAGCATTTTAT 467
1132 GAGATGGGGTTAATTAATGCTGCTG---GGGTCAAGAGCGAGTATGTTTCTGAGGT 1188
468 GATCAAGAAACAGCAGGAGACCTTGCGAAGATCGAGGGTCTCTTCTTAACTGGTGG 527
1189 GTTGGCAACTGATGCTCTCTTCAAGAGAGAGATTTGACATATTTCAATCAAC 1248
528 AAGAGCAATATATCGGTTCTCTAGCAGAAAGAGAACATGAGAGCTTTCATAGAC 587
1249 TCTGAGGCAAAACAGGCTGAATTTGAGTGAATCATACCAAGAGATGTTTGAAG 1308
588 TGCCAAAGGAAACCCGCTGAATATGAGATGATCTCTATATGAGATGATGAGGCC 647
QY 1309 GAGCTGAGGAGATCTCTGAGAGACATACGAGCTGAATCTTAAAGAAAGCTCTCA 1368

Db 648 CAGATGAACGAATATGAGAGATTAATCAACACTCAATTTCTGAGAAACAGGTGTT 707
QY 1369 AAAACAGAACAGCAGCCAGAGGCTGCTTGAAGATTTCTGGAATTTATGAGGAGAGCTG 1428
Db 708 AAGACAGACAGAGGCTCTAAAGTTTGAAGAAACCTGGGTGTTATTAACGAACTT 767
QY 1429 CGTAAGCTGACAGAGATTAATCGGATTCGAGACAGAGAACTAAGATCAGATGAACG 1488
Db 768 CGAGAGCTATGGAAGAAATATTTCTGAGAGTAAAGCTTAAGAGAACACTGAG 827
QY 1489 AACAGGAGAGATGATGATCAGACAGAGCTTTTCAATGATTTCAATCAACAGATTCAT 1548
Db 828 TAGAGAGAGAGATGAATATCCAGAAATTTCTTCATGATCAGATGAGAACTTCAC 887
QY 1549 GAAAGAGAGACCCAGAGAGAGAAATTTGAGATTTGACAGACAGAGAGAGCTGCAAG 1608
Db 888 AAGGCCACAG 947
QY 1609 GTTGTTCGACAGCAGCAGAGAACTTAATTCCTTGAAGATGACGATTCGGAAGAGA 1668
Db 948 GCT-----AGACATGTGATGATGATTTCTGGGACTAAGTGAATGCGAGGCTAAG 998
QY 1669 GCTGAGAGAGTGTCAAGCTTCATGAGTTTCAAGAGAAAGATGAGAGGTTGTGAA 1728
Db 999 AAGGACAGGTACAGCGGTCATAGAGTGCAGAGTTAAGACCTGAGAGGTTTGAAGCT 1058
QY 1729 GAGAGGAGATGCTGATTAAGATCAAGAGAGAGATGAGAGATGAAGAGAGGAT 1788
Db 1059 GAAAGGAGAGAAATGATTAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1118
QY 1789 CACGAGAGATATTTGATCTGAGAGAAAGATTTGATGAGCTTTGAGAACAGCTCATGTAC 1848
Db 1119 ATGCAAGAGAGGTGAGCTCGAAGAGAGTTGATGCTCTCACCGGTTGATGAG 1178
QY 1849 AAGCATGGCC 1858
Db 1179 AAGCATGAGC 1188
RESULT 12
AK061975 1214 bp mRNA linear PLN 24-JUL-2003
LOCUS Oryza sativa (japonica cultivar-group) cDNA clone:001-043-A02, full
DEFINITION Insert sequence.
ACCESSION AK061975
VERSION AK061975.1 GI:32971993
KEYWORDS FLI-CDNA; oligo-capping.
SOURCE Oryza sativa (japonica cultivar-group)
ORGANISM Oryza sativa (japonica cultivar-group)
REFERENCE The Rice Full-length cDNA Consortium, National Institute of
Agronomical Sciences Rice Full-length cDNA Project Team;
Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K.,
Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I.,
Kojima, K., Namiki, T., Ohneda, E., Yahagi, M., Sunuki, K., Li, C.,
Ohtsuki, K., Shishiki, T., Foundation of Advancement of International
Science Genome Sequencing & Analysis Group; Otsu, Y., Murakami, K.,
Iida, Y., Sugano, S., Fujimura, T., Suzuki, Y., Tsunoda, Y.,
Kurosaki, T., Kodama, T., Masuda, H., Kobayashi, M., Xie, Q., Lu, M.,
Narikawa, R., Sugiyama, A., Mizuno, K., Yokomizo, S., Nikura, J.,
Ikedo, R., Ishibiki, J., Kawamata, M., Yoshimura, A., Miura, J.,
Kusumegi, T., Oka, M., Ryu, R., Ueda, H., Aizawa, K., Arakawa, T., Fukuda, S.,
Kawai, J., Carninci, P., Adachi, J., Aizawa, K., Ishii, Y., Itoh, M.,
Hara, A., Hashidume, W., Hayatsu, N., Imotani, K., Ishii, Y., Ota, Y.,
Kagawa, I., Kondo, S., Kono, H., Miyazaki, A., Otsu, N., Ota, Y.,
Saito, R., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T.,
Yoshino, M., and Hayashizaki, Y.
Collection, mapping, and annotation of over 28,000 cDNA clones from
japonica rice

JOURNAL Science 301 (5631), 376-379 (2003)
 MEDLINE 22752273
 PUBMED 12869764
 REFERENCE
 AUTHORS

2 (bases 1 to 1214)
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Doi, K., Fujimura, T., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashizaki, Y., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Hotta, I., Iida, J., Iida, Y., Ikeda, R., Imamura, K., Imocani, K., Ishibiki, J., Ishii, Y., Ishikawa, M., Itoh, M., Kagawa, I., Kangawa, S., Katoh, H., Kawagashira, N., Kawai, J., Kawamata, M., Kikuchi, S., Kishikawa-Hirozane, T., Kishimoto, N., Kobayashi, M., Kodama, T., Kojima, K., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Kurosaki, T., Kusumegi, T., Li, C., Lu, M., Masuda, H., Matsubara, K., Matsuyama, T., Miura, J., Miyazaki, A., Mizuno, K., Murakami, K., Murata, M., Nagata, T., Nakamura, M., Namiki, T., Naitikawa, R., Nikura, J., Nishi, K., Nomura, K., Numasaki, R., Ohneda, E., Ohno, M., Ohtsuki, K., Oka, M., Ooka, H., Oeato, N., Ota, Y., Ootomo, Y., Ryu, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Satoh, K., Shibata, K., Shinagawa, A., Shiraki, T., Shishiki, T., Sogabe, Y., Sugano, S., Sugiyama, A., Suzuki, K., Suzuki, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Tsunoda, Y., Ueda, M., Waki, K., Xie, Q., Yahagi, W., Yamada, H., Yamamoto, M., Yasunishi, A., Yazaki, J., Yokomizo, S. and Yoshimura, A.

Direct Submission
 Submitted (05-DEC-2001) Shoshi Kikuchi, National Institute of Agricultural Sciences, Department of Molecular Genetics, Head of Laboratory of Gene Expression; 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan (E-mail: skikuchi@nias.affrc.go.jp, Tsukuba, Ibaraki Tel:81-29-838-7007, Fax:81-29-838-7007)
 This clone is one of the 28k full-length cDNA clones from japonica rice.

COMMENT

URL: http://cdna01.dna.affrc.go.jp/cDNA/
 NIAS Rice Full-Length cDNA Project Team: Kikuchi, S., Satoh, K., Nagata, T., Kawagashira, N., Doi, K., Kishimoto, N., Yazaki, J., Ishikawa, M., Yamada, H., Ooka, H., Hotta, I., Kojima, K., Namiki, T., Ohneda, E., Yahagi, W., Suzuki, K., Li, C., Ohtsuki, K., Shishiki, T. and Yamamoto, M.

PAIS Genome Sequencing & Analysis Group: Ootomo, Y., Iida, Y., Fujimura, T., Ikeda, R., Ishibiki, J., Kawamata, M., Kobayashi, M., Kodama, T., Kurosaki, T., Kusumegi, T., Lu, M., Masuda, H., Miura, J., Mizuno, K., Naitikawa, R., Nikura, J., Oka, M., Ryu, R., Sugano, S., Sugiyama, A., Suzuki, Y., Tsunoda, Y., Ueda, M., Xie, Q., Yokomizo, S., Yoshimura, A., Matsubara, K. and Murakami, K.
 Genomic Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken: Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P., Fukuda, S., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hori, F., Iida, J., Imamura, K., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oeato, N., Ota, Y., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES

Location/Qualifiers
 1..1214
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultiVar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="001-043-A02"

ORIGIN

Query Match 15.4%; Score 288.8; DB 8; Length 1214;
 Best Local Similarity 58.7%; Pred. No. 2e-59;
 Matches 544; Conservative 0; Mismatches 367; Indels 15; Gaps 2;
 QY 934 ACTAGACTGATTAAGACGATTAAGATGCTCGGATGGCAACCAAGACTGCTG 933

Db 2 ACTGATTTGGAACAAGATAGAGATGATTAAGTGGAAAGGATATGGGGAAACCAAGAGCTGAATT 61
 QY 994 GAATCTTCGACAAAGATAGAGCTCTTAGAGACCGCATCTCTATAGTGTCCACAGGCCCAT 1053
 Db 62 GATTATTTCAGTAATATCTCCAGCAAGCAAGCAGCATGCTTATGTCTCAATATGGGAC 121
 QY 1054 CATTGAGATAGTCTCTGATGTTTGGAGACAGTGCACATGCTATTTGGAGGCCCAAGC 1113
 Db 122 CATTGATAGATGTGTATCTTTGACAGCTGGCTGTGGCTATATGAGAGAGACT 181
 QY 1114 CTCCACCGGAGCTTAGCTGAGATGGGTTAGATGAATTCCTGG-----GCTGGAAG 1167
 Db 182 TTGACATGATCATCTTTGTTGCTGATCAAGAAACAGACAGGAATATATGAAACAGTGTCTCAAG 241
 QY 1168 CCGAGATGTTTTTGGAGAGTGTGGCCCACTGTATGCTTCTTCCACAGGAAGAAAT 1227
 Db 242 GTTACATTTCTTACCTGTGTGGAAAGGACGCTGTATGTCTTATGCTTCAAAAGATGAC 301
 QY 1228 CTGACATATTCATCACTCTCAAGGCAAAACAAAGGCTGAATTCAGTTGAATCA 1287
 Db 302 ATGAGACATTTCAATAGACATTTGCCAGGAAATCCCTTTGAATATAGATGATCT 361
 QY 1288 TACCAAGAGATGTTGTAAAGAGCTGAGCAGATCTTGAGACAAATGAGCAGCTGAAC 1347
 Db 362 TATTAATGATGATGTGTATCCCAATGAAGCAGATGAGGAAACAAACAGCACTGAAT 421
 QY 1348 TACTTTAAGAACAGCTCTCAAAACAGAACAGCAAGCCCAAGTCTTGAAGAAATCTCG 1407
 Db 422 TACCTGAAACAAATATGTTTAAAGAGCAGCACTTAAATTAAGATGAAGTAACTATTG 481
 QY 1408 GAAATTAAGAGGAGAGCTGCTAGAACTGCAGAGATTAATCGATCTGTAGACAGAGA 1467
 Db 482 AGTGTGTTTATCCAGAACTTAGGAGACAAATGGAAGAAATATCTATTGATGAATTA 541
 QY 1468 ACTAAGATCAGCATGAACAGAACAGGAGAGATGATGACACAGCAGGTTTTTCATG 1527
 Db 542 GCCAAAGAAACACATGAGATGAGAAAGATGAATGAATCAACAGAGAAATTTTCCAT 601
 QY 1528 GATTCATTAACAGATTCATGAAGAGAGACGCAAGAGAGAAATTTCCAGATCTTG 1587
 Db 602 GATCAATGAAATTAATTCACAGGCAACAGAGAGAAAGAAATCAAGTTTGAGAGTTG 661
 QY 1588 CAGCAGCAGAAACGTCGCAAGGTTGTGGCCAGCAGCAGACAGAACTTAATCCCTTAG 1647
 Db 662 CTGCAAGAGAGAGCAAAAGCT-----AGACAAAGATGATTTGATCTTGAAGT 712
 QY 1648 AATGACGATTTGCCAAGAGAGCTGAGAGATGTCAGCTTCATCGATTTCAAGAGAA 1707
 Db 713 ACTGAAGATCGCAGGCAAGAGAGAAAGAAAGATTCAAACTTCATGATGACCAAGTGAAG 772
 QY 1708 GAGATGAGAGATTTGTGAGAGAGAGAGAGATGCTGATTAAGATGAAGAAAGATG 1767
 Db 773 GATGTTGAGAAATTTGACCCCAAGCAGACAACTGATCAAAATTTGATGAAGAGAAAG 832
 QY 1768 GAAGACATGAAGAGGATCAGCAGAGATTTGATCTGAGAGAAAGATTTGATGAG 1827
 Db 833 GTGAGCTCAAGAGAGATGATCTGGGAGAGAAATTTGATCTGAGAGAGAGCTGACACA 892
 QY 1828 GCTTTGGAACAGCTCATGTACAGCA 1853
 Db 893 GCCCTCATCTCCCTGATGATGAACA 918

RESULT 13
 AF469493
 LOCUS AF469493 513 bp mRNA linear PLN 02-JUL-2003
 DEFINITION Triticum aestivum H07R mRNA, partial cds.
 ACCESSION AF469493
 VERSION AF469493.1 GI:3240075
 KEYWORDS
 SOURCE Triticum aestivum (bread wheat)
 ORGANISM Triticum aestivum

REFERENCE
AUTHORS
JOURNAL
TITLE
SUBMITTED (18-JAN-2002) College of Life Science, Shandong Agricultural University, Dai zong Street 61, Taiman, Shandong 271018, P.R. China

FEATURES
source

CDS

/note="similar to HOTA of Hordeum vulgare"
/protein_id="AAP0610.1"
/db_xref="GI:32400756"
/translation="MILGMNOLLEVEPSDYATKRAHAYGPGHGMVLIPESSAVG
YMEERLHKHFDKTRDTRDNRVPRFPGGRKOLYGFARDEFTENRBCOGKSR
LKYEMRSHNEVVAQKQMSQEDNOQNLNKLKVKTKKESKVEETLGVITOKLHETM
ENIFVRSKA"

ORIGIN

Query Match 10.2%; Score 192; DB 8; Length 513;
Best Local Similarity 62.3%; Pred. No. 8.5e-36;
Matches 319; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

964 TGGCTCGGATGGGCAACCAAGAGCTGTGGAATCTTCGACAGTNGAGGCTTAGA 1023
1 TGGTTAGGATGGGCAACCAAGAGCTGTGGAATCTTCGACAGTNGAGGCTTAGA 60
1024 GCAAGCCATTCCTATGCTCCAGAGGCAATGCGGATGATGCTTCTGATGTTAGAGC 1083
61 GCAAGCCATTCCTATGCTCCAGAGGCAATGCGGATGATGCTTCTGATGTTAGAGC 120
1084 AGTGCACCTGCTATTTGGAGCCGAGCGCTTCCACCGGAGTTAGCTGAGATGGGCTTA 1143
121 TCTGCTGTGGGCTATTTGGAGCCGAGCGCTTCCACCGGAGTTAGCTGAGATGGGCTTA 180
1144 GATGAAATTCCTG--GGGTGAGAGCGCAGATGCTTCTGAGAGTTCGCCACTG 1200
181 GACAGGGACCTTGGGCAATGCGGATGATGCTTCTGAGAGTTCGCCACTG 240
1201 TATGGCTTCTTCCAGAGCAAGATCTGACATTAATCAATCACTCTCAAGGCAAA 1260
241 TACGGTTTCTTCCAGAGCAAGATCTGACATTAATCAATCACTCTCAAGGCAAA 300
1261 ACAAGGCTGAATTCGAGTTGAATCTATACCAAGAGATGTTTAAAGAGCTGAGGCG 1320
301 AGCGGCTGAATTCGAGTTGAATCTATACCAAGAGATGTTTAAAGAGCTGAGGCG 360
1321 ATCTCTGAGGACATTCAGAGCTGATCTTAAAGAACTTCAAAAACAGAAACAG 1380
361 ATGAGCGAAGATTAATCAACTCAATTAATCTGAGAAACAGAGGTTTAAAGAGGCG 420
1381 CACGCGAAGTCTGAGAGATCTCTGGAATTTAGAGCGAAGAGCTGAGAGCTGCA 1440
421 CGCTTAAGTGTGAGAGAAACCTCGGTGTATTTAGCGAGAACTTCAAGAGACTTAA 480
1441 GAGGTAATCGATGCTGAGAGCAAGACTTA 1472
481 GAAAGAAATATTCGTCTCAGAGTAAGCTTA 512

RESULT 14
CNS08CNO/c 91053 bp DNA linear PLN 21-NOV-2003
LOCUS Oryza sativa chromosome 12, BAC OJ1561_A05 of library Monsanto
DEFINITION from chromosome 12 of cultivar Nipponbare of ssp. japonica of Oryza
sativa (rice), complete sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

Center code: GS
Web site: http://www.genoscope.cns.fr/
Contact: Segre@genoscope.cns.fr

The following sequence is oriented from the T7 to the SP6 end. The
nucleotide sequence of this BAC clone was generated by combining
Monsanto, Syngenta and Genoscope sequencing data.
Upstream BAC (overlapping the T7 end) : OSUN0029N15 (AC-BX536967)
Downstream BAC (overlapping the SP6 end) : OJ1102_B11 (AC-AL713901)
FINISHED SEGMENT STARTS AT BASE 1
FINISHED SEGMENT ENDS AT BASE 75085

FEATURES

source
Location/Qualifiers
1..91053
/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="genomic DNA"
/cultivar="Nipponbare"
/sub_species="japonica"
/db_xref="taxon:39947"
/chromosome="12"
/clone="OJ1561_A05"
/clone_1ib="Monsanto"

ORIGIN

Query Match 9.0%; Score 169.4; DB 8; Length 91053;
Best Local Similarity 61.1%; Pred. No. 2.6e-30;
Matches 310; Conservative 0; Mismatches 191; Indels 6; Gaps 2;

460 TTTCCTGAGCTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 519
81521 TTGCTGCTGCTGATCTCTGAAACAGCAATGCTGAAAGTTAGACATCGAGCATGAT 81462
520 TTGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 579
81461 AATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 81402
580 AGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 639
81401 AGTTTGAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 81342
640 TTGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 699
81341 TTAAGTGAAGACGTTAATGAAACCAAGAGGATGATGATGATGATGATGATGATGAT 81282

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 14:24:13 ; Search time 494.227 Seconds
(without alignments)
16142.601 Million cell updates/sec

Title: US-10-030-829-2

Perfect score: 1878
Sequence: 1 atgagcttcaggcgctgcctcc.....ttcacatgaagatgatga 1878

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapept 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: geneseqn1808:*
2: geneseqn1908:*
3: geneseqn2008:*
4: geneseqn2001as:*
5: geneseqn2001bs:*
6: geneseqn2002as:*
7: geneseqn2003as:*
8: geneseqn2003bs:*
9: geneseqn2003cs:*
10: geneseqn2004as:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1878	100.0	1878	4	AAFP25374 Nucleotide
2	1462	77.8	3275	4	AAFP25373 Nucleotide
3	75	4.0	272	7	ABX81661 Corn ear-
4	74.4	4.0	274	7	ABX82003 Corn ear-
5	63.6	3.4	310	6	ABL171922 Corn tass
6	58.8	3.1	3489	3	AAA30290 Kapori's
7	58.8	3.1	3489	4	AAFP2901 Nucleotide
8	58.8	3.1	3489	6	ABA33467 Nucleotide
9	58.8	3.1	32207	2	AAV73805 KSHV LUR
10	58.8	3.1	137507	2	AAV19941 KSHV long
11	56.2	3.0	1080	6	AB199537 Mouse isc
12	55	2.9	1686	2	AAO87587 DNA encod
13	55	2.9	4000	2	AAO87587 DNA encod
14	53.2	2.8	1140	3	AAO87587 DNA encod
15	53.2	2.8	16442	2	AAO87587 DNA encod
16	53	2.8	3399	2	AAO87587 DNA encod
17	52.6	2.8	700	4	ABA65105 Human foe
18	52.6	2.8	700	4	ABA65105 Human foe
19	52.6	2.8	700	4	ABA65105 Human foe
20	52.6	2.8	700	4	ABA65105 Human foe
21	52.6	2.8	700	4	ABA65105 Human foe
22	52.6	2.8	700	4	ABA65105 Human foe
23	52.6	2.8	700	4	ABA65105 Human foe

24	52.6	2.8	700	4	AB838852 Human liv
25	52.6	2.8	700	5	AA105782 Human #57
26	52.6	2.8	700	6	AA105782 Human #57
27	52.2	2.8	2000	7	ADA71938 Rice gene
28	52	2.8	766	4	AA122114 Probe #12
29	52	2.8	766	4	ABA67193 Human foe
30	52	2.8	766	4	ABA67193 Human foe
31	52	2.8	766	4	ABA67193 Human foe
32	52	2.8	766	4	ABA67193 Human foe
33	52	2.8	766	4	ABA67193 Human foe
34	52	2.8	766	4	ABA67193 Human foe
35	52	2.8	766	4	ABA67193 Human foe
36	52	2.8	766	4	ABA67193 Human foe
37	52	2.8	766	4	ABA67193 Human foe
38	52	2.8	766	4	ABA67193 Human foe
39	52	2.8	766	4	ABA67193 Human foe
40	51.6	2.7	2093	9	AD863512 Human CDN
41	51.2	2.7	693	5	AA874240 DNA encod
42	51.2	2.7	693	5	AA874240 DNA encod
43	51	2.7	267	4	AAK45604 Human bon
44	51	2.7	267	4	AAK45604 Human bon
45	51	2.7	267	4	AAK45604 Human bon

ALIGNMENTS

RESULT 1
ID AAFP25374 standard; cDNA; 1878 BP.
XX AAFP25374;
AC AAFP25374;
DT 15-MAY-2001 (first entry)
XX Nucleotide sequence of the Arabidopsis SGS3 polypeptide.
DE SGS3 gene; post-transcriptional inactivation; RNA degradation;
XX viral resistance; resistance; fatty acid content; protein content; ss.
KM Arabidopsis thaliana.
OS Arabidopsis thaliana.
XX Key Location/Qualifiers
FH 1.1878
FT CDS /tag= a
FT /product= "SGS3"
XX MO200105951-A2.
XX 25-JAN-2001.
XX 13-JUL-2000; 2000MO-FR002052.
XX 16-JUL-1999; 99FR-00009417.
XX 26-JAN-2000; 2000FR-00001006.
XX (AVET) AVENTIS CROSCIENCE SA.
XX (INRG) INST NAT RECH AGRONOMIQUE.
XX Beclin C, Elmeyan T, Vaucheret H;
XX WPI; 2001-159529/16.
XX P-RSDB; AAB31798.
XX New SGS3 gene from Arabidopsis thaliana, useful for increasing virus
XX resistance in plants and, when inhibited, for increasing transgene
XX expression.
XX Claim 1; Page 32-35; 36pp; French.
XX The present sequence encodes the Arabidopsis thaliana SGS3 polypeptide.
XX The SGS3 gene is essential for post-transcriptional inactivation
XX (degradation of RNA) and for resistance to viruses. Overexpression of

CC SSS3 results in plants with increased resistance to viruses, while
 CC inactivation of SSS3 in transgenic plants (e.g. by expressing an antisense
 CC RNA, by mutation or by homologous recombination) increases the level of
 CC the transgene product. This product may e.g. impart resistance (to
 CC herbicide, insects or pathogens), alter contents of essential fatty acids
 CC or proteins, or is pharmaceutically active, e.g. an immunoglobulin or
 CC interferon

XX Sequence 1878 BP; 575 A; 332 C; 554 G; 417 T; 0 U; 0 Other;

Query Match 100.0%; Score 1878; DB 4; Length 1878;

Best Local Similarity 100.0%; Pred. No. 0; Mismatches 0; Indels 0; Gaps 0;

Matches 1878; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAGTTCTAGGCTGCTCCATGCTCAAGTAAAGAAAGAAAGTTCAGGAGGTTATAGGCT 60
 1 ATGAGTTCTAGGCTGCTCCATGCTCAAGTAAAGAAAGAAAGTTCAGGAGGTTATAGGCT 60
 61 GAGGTGAACAGTTGTTCAAGTTGGCAGGAGCAGACTGGCTTCTCAAGATGAT 120
 61 GAGGTGAACAGTTGTTCAAGTTGGCAGGAGCAGACTGGCTTCTCAAGATGAT 120
 121 GGAGAGAGTGGAGGTCATTTCCAGAGAAACAGAAACAGAAACAGAAACAGTTCTGGA 180
 121 GGAGAGAGTGGAGGTCATTTCCAGAGAAACAGAAACAGAAACAGAAACAGTTCTGGA 180
 121 GGAGAGAGTGGAGGTCATTTCCAGAGAAACAGAAACAGAAACAGAAACAGTTCTGGA 180
 181 AAAAATTGGGTTCTCAAGATTCCTCTAGAGCTTGGGTTGTCAGACAGAGG 240
 181 AAAAATTGGGTTCTCAAGATTCCTCTAGAGCTTGGGTTGTCAGACAGAGG 240
 181 AAAAATTGGGTTCTCAAGATTCCTCTAGAGCTTGGGTTGTCAGACAGAGG 240
 241 AGAGTACCAACGTAATCTGGAGAGAGAAACATGATTCGGAGAGAGTAAAGCAATGCT 300
 241 AGAGTACCAACGTAATCTGGAGAGAGAAACATGATTCGGAGAGAGTAAAGCAATGCT 300
 241 AGAGTACCAACGTAATCTGGAGAGAGAAACATGATTCGGAGAGAGTAAAGCAATGCT 300
 301 CGGGGATTCAGAGCTTAACATATCTGGTGGGAGCAGGCTTGAACAGAAATGATGATAC 360
 301 CGGGGATTCAGAGCTTAACATATCTGGTGGGAGCAGGCTTGAACAGAAATGATGATAC 360
 301 CGGGGATTCAGAGCTTAACATATCTGGTGGGAGCAGGCTTGAACAGAAATGATGATAC 360
 361 AACTTTGTGGACCCCACTGTATCTGCGCTCTTTGGAGAGAGATGGAATTTGGCAG 420
 361 AACTTTGTGGACCCCACTGTATCTGCGCTCTTTGGAGAGAGATGGAATTTGGCAG 420
 361 AACTTTGTGGACCCCACTGTATCTGCGCTCTTTGGAGAGAGATGGAATTTGGCAG 420
 421 GCAAGAGAGGTTCTGCTCAAGACAGCTGTCAGAGTTTCTGAGTGAAGATGAT 480
 421 GCAAGAGAGGTTCTGCTCAAGACAGCTGTCAGAGTTTCTGAGTGAAGATGAT 480
 421 GCAAGAGAGGTTCTGCTCAAGACAGCTGTCAGAGTTTCTGAGTGAAGATGAT 480
 481 GTCGATTAATCTCTGAGAGAGATGATTCGATCTTTGATGATTCGATGATGATGAT 540
 481 GTCGATTAATCTCTGAGAGAGATGATTCGATCTTTGATGATTCGATGATGATGAT 540
 481 GTCGATTAATCTCTGAGAGAGATGATTCGATCTTTGATGATTCGATGATGATGAT 540
 541 CTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 541 CTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 541 CTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 541 CTTCGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 600
 601 AATTAAGGTTCAAAAAGTTCTTTGGCAGCTTGGATGCTTGTGATGAGATGATGAT 660
 601 AATTAAGGTTCAAAAAGTTCTTTGGCAGCTTGGATGCTTGTGATGAGATGATGAT 660
 601 AATTAAGGTTCAAAAAGTTCTTTGGCAGCTTGGATGCTTGTGATGAGATGATGAT 660
 601 AATTAAGGTTCAAAAAGTTCTTTGGCAGCTTGGATGCTTGTGATGAGATGATGAT 660
 661 GAACCAAGAGAGGCTGTCATGTCAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 720
 661 GAACCAAGAGAGGCTGTCATGTCAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 720
 661 GAACCAAGAGAGGCTGTCATGTCAGCTTGTGAGAGAGAGAGAGAGAGAGAGAGAG 720
 721 TATTAAGTGAACCTCTACTAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 721 TATTAAGTGAACCTCTACTAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 721 TATTAAGTGAACCTCTACTAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 721 TATTAAGTGAACCTCTACTAGCTCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 780
 781 CATAGAGATTTGGCTGAAGTTTGAAGAGATCTAAGAGAGAGAGAGAGAGAGAGAG 840
 781 CATAGAGATTTGGCTGAAGTTTGAAGAGATCTAAGAGAGAGAGAGAGAGAGAGAG 840
 781 CATAGAGATTTGGCTGAAGTTTGAAGAGATCTAAGAGAGAGAGAGAGAGAGAGAG 840
 781 CATAGAGATTTGGCTGAAGTTTGAAGAGATCTAAGAGAGAGAGAGAGAGAGAGAG 840
 841 CCTGTGTGAGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900
 841 CCTGTGTGAGATTTATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 900

901 ATTGTGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 901 ATTGTGCTCTCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 960
 961 AATGTGCTGAGATGAG 1020
 961 AATGTGCTGAGATGAG 1020
 961 AATGTGCTGAGATGAG 1020
 961 AATGTGCTGAGATGAG 1020
 1021 AG 1080
 1021 AG 1080
 1021 AG 1080
 1021 AG 1080
 1081 AG 1140
 1081 AG 1140
 1081 AG 1140
 1081 AG 1140
 1141 TTAGATTAAGATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1141 TTAGATTAAGATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1141 TTAGATTAAGATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1141 TTAGATTAAGATTCCTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1200
 1201 TATGCTCTCTGAG 1260
 1201 TATGCTCTCTGAG 1260
 1201 TATGCTCTCTGAG 1260
 1201 TATGCTCTCTGAG 1260
 1261 ACAAGGCTGAATTCGAGTGAATTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 1261 ACAAGGCTGAATTCGAGTGAATTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 1261 ACAAGGCTGAATTCGAGTGAATTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 1261 ACAAGGCTGAATTCGAGTGAATTCATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1320
 1321 ATCTCTGAG 1380
 1321 ATCTCTGAG 1380
 1321 ATCTCTGAG 1380
 1321 ATCTCTGAG 1380
 1381 CACGCAAGTGTCTGAG 1440
 1381 CACGCAAGTGTCTGAG 1440
 1381 CACGCAAGTGTCTGAG 1440
 1381 CACGCAAGTGTCTGAG 1440
 1441 GAGATTAATTCGATGAG 1500
 1441 GAGATTAATTCGATGAG 1500
 1441 GAGATTAATTCGATGAG 1500
 1441 GAGATTAATTCGATGAG 1500
 1501 ATGATGAG 1560
 1501 ATGATGAG 1560
 1501 ATGATGAG 1560
 1501 ATGATGAG 1560
 1561 GCAAG 1620
 1561 GCAAG 1620
 1561 GCAAG 1620
 1561 GCAAG 1620
 1621 CAGCAG 1680
 1621 CAGCAG 1680
 1621 CAGCAG 1680
 1621 CAGCAG 1680
 1681 TCAAGCTTCATGAG 1740
 1681 TCAAGCTTCATGAG 1740
 1681 TCAAGCTTCATGAG 1740
 1681 TCAAGCTTCATGAG 1740
 1741 CTGATTAAGATTAAG 1800
 1741 CTGATTAAGATTAAG 1800
 1741 CTGATTAAGATTAAG 1800
 1741 CTGATTAAGATTAAG 1800
 1801 TTTGATCTGAG 1860
 1801 TTTGATCTGAG 1860
 1801 TTTGATCTGAG 1860
 1801 TTTGATCTGAG 1860
 1861 CACAATGAAGATGATGA 1878
 1861 CACAATGAAGATGATGA 1878
 1861 CACAATGAAGATGATGA 1878
 1861 CACAATGAAGATGATGA 1878
 1861 CACAATGAAGATGATGA 1878

RESULT 2
 AAF25373
 ID AAF25373 standard; DNA; 3275 BP.

AC AAF25373;
 XX
 DT 15-MAY-2001 (first entry)
 XX
 DE Genomic sequence of the Arabidopsis SGS3 gene.
 XX
 KW SGS3 gene; post-transcriptional inactivation; RNA degradation;
 KM vital resistance; resistance; fatty acid content; protein content; ss.
 XX
 OS Arabidopsis thaliana.
 XX
 FH Key Location/Qualifiers
 FT exon 666..1658
 FT /*tag= a
 FT intron 1659..1731
 FT /*tag= b
 FT exon 1732..2023
 FT /*tag= c
 FT intron 2024..2134
 FT /*tag= d
 FT exon 2135..2379
 FT /*tag= e
 FT intron 2380..2481
 FT /*tag= f
 FT exon 2482..2648
 FT /*tag= g
 FT intron 2649..2738
 FT /*tag= h
 FT exon 2739..2949
 FT /*tag= i
 FT
 FN WO200105951-A2.
 PN
 PD 25-JAN-2001.
 PP 13-JUL-2000; 2000MO-FR002052.
 PR 16-JUL-1999; 99FR-00009417.
 PR 26-JAN-2000; 2000FR-00001006.
 XX
 PA (AVET) AVENTIS CROPS SCIENCE SA.
 PA (INRG) INST NAT RECH AGRONOMIQUE.
 XX
 PI Beclin C, Elmayer T, Vaucheret H;
 DR MPI; 2001-159529/16.
 XX
 PT New SGS3 gene from Arabidopsis thaliana, useful for increasing virus
 PT resistance in plants and, when inhibited, for increasing transgene
 PT expression.
 XX
 PS Claim 1; Page 31-32; 36pp; French.
 XX
 CC The present sequence represents the genomic sequence of the Arabidopsis
 CC thaliana SGS3 gene. The SGS3 gene is essential for post-transcriptional
 CC inactivation (degradation of RNA) and for resistance to viruses.
 CC Overexpression of SGS3 results in plants with increased resistance to
 CC viruses, while inactivation of SGS3 in transgenic plants (e.g. by
 CC expressing antisense RNA, by mutation or by homologous recombination)
 CC increases the level of the transgene product. This product may e.g.
 CC impart resistance (to herbicide, insects or pathogens), alter contents of
 CC essential fatty acids or proteins, or is pharmaceutically active, e.g. an
 CC immunoglobulin or interferon
 SO Sequence 3275 BP, 956 A; 561 C; 804 G; 954 T; 0 U; 0 Other;

Db	696	ATGAGTCTACAGGCTGGTCCAAATGCTTAAGAAAAAGAAACGTTACAGGGTGGTTAATGAGCT	755
QY	61	GAGGTGGAACGTTGCTTCAAGGTTTGGCAGAGGACGAGACTGCTTCTTCAACAAGATGAT	120
Db	756	GAGGTGGAACGTTGGTTCAGAGTTTGGCAGAGGACGAGACTGCTTCTTCAACAAGATGAT	815
QY	121	GGAGGAGGTGGGAGGTCAATTTCCAAAGAAACAAGAAACAACAGGAAAACACTTCTGGA	180
Db	816	GGAGGAGGTGGGAGGTCAATTTCCAAAGAAACAAGAAACAACAGGAAAACACTTCTGGA	875
QY	181	AAAACCTTGGTTCAGAAATTCGAATCTTCTAGAGCTTGGGGTGTCCAGACAAAGG	240
Db	876	AAAACCTTGGTTCAGAAATTCGAATCTTCTAGAGCTTGGGGTGTCCAGACAAAGG	935
QY	241	AGAGGTACCAACGTATCTGGGAGAGGAAACAATGATCTGGGAGAGGTAAACGGCAATGCT	300
Db	936	AGAGGTAGCAACGTATCTGGGAGAGGAAACAATGATCTGGGAGAGGTAAACGGCAATGCT	995
QY	301	CGGGGCATTCAAGCTAACATATCTGAGTCCGGGGAAGAGCGTTGAGACAGAAATGATGATAC	360
Db	996	CGGGGCATTCAAGCTAACATATCTGAGTCCGGGGAAGAGCGTTGAGACAGAAATGATGATAC	1055
QY	361	AACCTTGGGCAACCCCAACCTGATCTCGCCCTCTTCTGGAAGAGGTGAATTTGGCAG	420
Db	1056	AACCTTGGGCAACCCCAACCTGATCTCGCCCTCTTCTGGAAGAGGTGAATTTGGCAG	1115
QY	421	GCAAGAGAGGTTCTGCTCAGACACACAGCTGTGAGAGAGTTTCTGACGTGAGAGATGAT	480
Db	1116	GCAAGAGAGGTTCTGCTCAGACACACAGCTGTGAGAGAGTTTCTGACGTGAGAGATGAT	1175
QY	481	GTGGATATGCTTCTCGAAGAAAGAAATGATTTCCGATGCTTTGGATGATTTCTGATGACAC	540
Db	1176	GTGGATATGCTTCTCGAAGAAAGAAATGATTTCCGATGCTTTGGATGATTTCTGATGACAC	1235
QY	541	CTTGCAGGTGATGATTTGACTCGGATGTGATGATGATGATGATGATGATGATGATGATGAT	600
Db	1236	CTTGCAGGTGATGATTTGACTCGGATGTGATGATGATGATGATGATGATGATGATGATGAT	1295
QY	601	AATAAGTGTTCAAAAAGTTCTTGTGGCAGCTTGGAGTATGCTTGTGATCGAGCAGATTAAT	660
Db	1296	AATAAGTGTTCAAAAAGTTCTTGTGGCAGCTTGGAGTATGCTTGTGATCGAGCAGATTAAT	1355
QY	661	GAACCAACAGAGGCACTGGCATTGTCCAGCTTGTGAGAACGAGCCTGGTGCATGATGTTGG	720
Db	1356	GAACCAACAGAGGCACTGGCATTGTCCAGCTTGTGAGAACGAGCCTGGTGCATGATGTTGG	1415
QY	721	TATTAACCTGCAACCTCTATACGTCATGCGACGACAAAGAGAGCTAGCCAGTTAAGCTC	780
Db	1416	TATTAACCTGCAACCTCTATACGTCATGCGACGACAAAGAGAGCTAGCCAGTTAAGCTC	1475
QY	781	CATAGAGAAATGCGCTGAGAAATTTTGAAGAAAGGATCTACAGATGAGAGCGCATCTGTCAAT	840
Db	1476	CATAGAGAAATGCGCTGAGAAATTTTGAAGAAAGGATCTACAGATGAGAGCGCATCTGTCAAT	1535
QY	841	CCTTGTGGTGAATTTATGAGCAGTGGAAAGGTTGGGTGAGATGAAGAAAGATTTATGAA	900
Db	1536	CCTTGTGGTGAATTTATGAGCAGTGGAAAGGTTGGGTGAGATGAAGAAAGATTTATGAA	1595
QY	901	ATTGTCTGCGCTCCAAATGTGTCATCATGTAATCTAGACTGAGTAAAGACATGAATGAT	960
Db	1596	ATTGTCTGCGCTCCAAATGTGTCATCATGTAATCTAGACTGAGTAAAGACATGAATGAT	1655
QY	961	A-----	961
Db	1656	AAGGTGAATCTTCTGCTTTTACTTTTAAATTTTCTTCTGCAATTTCTAAGATCTTA	1715
QY	962	-----AGTGGCTCGGCATGGGCAACCAAGAGCTGTGGAATATCTTGCACAA	1007
Db	1716	GAATGTTAACATGTAAGTGGCTCGGCATGGGCAACCAAGAGCTGTGGAATATCTTGCACAA	1775
QY	1008	GTATGAGGCTTTTAGACACGCCATTTCTATATGTGTCCACAGGGCCATGTGTGGAATGATGT	1067
Db	1776	GTATGAGGCTTTTAGACACGCCATTTCTATATGTGTCCACAGGGCCATGTGTGGAATGATGT	1835


```

QY 1068 TCTGATGTTGAGAGAGAGTGCACATGCTGATTTTGAGGCGGAAACGCTCCACCGGGAGTT 1127
DB 1836 TCTGATGTTGAGAGAGAGTGCACATGCTGATTTTGAGGCGGAAACGCTCCACCGGGAGTT 1895
QY 1128 AGCTGAGATGGGGTTAGATAGAAATTCCTGGGGTCAAGAACCGCATATGTTTCTGAGG 1187
DB 1896 AGCTGAGATGGGGTTAGATAGAAATTCCTGGGGTCAAGAACCGCATATGTTTCTGAGG 1955
QY 1188 TGTGGCCAACTGTATGAGCTTCTTCCAAAGAGCAAGATCTGAGACATATTCATCAACA 1247
DB 1956 TGTGGCCAACTGTATGAGCTTCTTCCAAAGAGCAAGATCTGAGACATATTCATCAACA 2015
QY 1248 CTCTCA----- 1253
DB 2016 CTCTCAAGTTCTCTCCCAAGAAATTTGATATATGCTTTTATGTTTGTCAATTGGAAT 2075
QY 1254 -----AGG 1256
DB 2076 TTAAGTTTGTGTCGGTAAATGATCATCTGTATATATATCTATGATTCAATTAGG 2135
QY 1257 CAAAACAGGCTGAAATTTGAGTGAATCATACCAAGATGTTTAAAGAGCTGAG 1316
DB 2136 CAAAACAGGCTGAAATTTGAGTGAATCATACCAAGATGTTTAAAGAGCTGAG 2195
QY 1317 GCAGATCTCTGAGAGCAATCAGACGCTGAACCTTTAAGAACAGCTCTCAAAAACAGAA 1376
DB 2196 GCAGATCTCTGAGAGCAATCAGACGCTGAACCTTTAAGAACAGCTCTCAAAAACAGAA 2255
QY 1377 CAAAGCAGCAGAGTCTTGAAGAAATCTGTGAATTTATGAGGAGAACCTGCTGTAAC 1436
DB 2256 CAAAGCAGCAGAGTCTTGAAGAAATCTGTGAATTTATGAGGAGAACCTGCTGTAAC 2315
QY 1437 TCAGAGGATATTCGATGCTGAGAGCAAGAACTAAGATGCGCAGAAACAGAGAGGA 1496
DB 2316 TCAGAGGATATTCGATGCTGAGAGCAAGAACTAAGATGCGCAGAAACAGAGAGGA 2375
QY 1497 AG----- 1498
DB 2376 AAGAGTATGATTTTCTGAGAAATCAAAACCTTGAATTTGATTAATCACTGATTTCA 2435
QY 1499 -----AGATGATGCAACGA 1514
DB 2436 CATTTTGTATATTTGTCACAAACAAAAAAGCTGTGCTGATTAAGATGATGATGCAACGA 2495
QY 1515 CAGGTTTTCATGATGATTTCAATCAAAAGATCCATGAAAAGAGAGCGCAAGAGAGAA 1574
DB 2496 CAGGTTTTCATGATGATTTCAATCAAAAGATCCATGAAAAGAGAGCGCAAGAGAGAA 2555
QY 1575 TTTTCGATGTTGCACAGACAGAAAGTGCAGAGTTGTTGGCCAGCAGCAGAGCAACAT 1634
DB 2556 TTTTCGATGTTGCACAGACAGAAAGTGCAGAGTTGTTGGCCAGCAGCAGAGCAACAT 2615
QY 1635 TAATCCCTTAGCAATGACGATGCGCAAG----- 1665
DB 2616 TAATCCCTTAGCAATGACGATGCGCAAG----- 2675
QY 1666 ----- 1665
DB 2676 CTGGCGTTTGTTTTCAAACTAGAGTAATCTGAATTAATCCGTTTGTGATTTCTTTCG 2735
QY 1666 -AGAGCTGAGAGAGTGTCAAGCTTTCAGAGTTTCAAGAGAAAGAGATGAGAGTTTGT 1724
DB 2736 CAGAGCTGAGAGAGTGTCAAGCTTTCAGAGTTTCAAGAGAAAGAGATGAGAGTTTGT 2795
QY 1725 GGAAGAGAGAGAGATCTGATTAAGATCAAGAGAAAGATGAGAGATCAAGAGAGAG 1784
DB 2796 GGAAGAGAGAGAGATCTGATTAAGATCAAGAGAAAGATGAGAGATCAAGAGAGAG 2855
QY 1785 GCATCAGAGAGATTTTGTATGAGAGAAAGATTTGATGAGGCTTTGGAACAGCTCAT 1844
DB 2856 GCATCAGAGAGATTTTGTATGAGAGAAAGATTTGATGAGGCTTTGGAACAGCTCAT 2915

```

```

QY 1845 GTACAAGCATGGCCTTCAACATGAAGATGATTGA 1878
DB 2916 GTACAAGCATGGCCTTCAACATGAAGATGATTGA 2949

RESULT 3
ABX81661
ID ABX81661 standard; cDNA; 272 BP.
XX
AC ABX81661;
XX
DT 24-APR-2003 (first entry)
XX
DE Corn ear-derived polynucleotide (cpd) #121.
XX
KW Corn ear-derived polynucleotide; cdp; cDNA library; SATMON022; SATMON023;
KW structural gene; functional gene; regulatory gene;
KW corn ear-specific profile; gene transcription; gene expression;
KW hybrid plant; desirable trait expression; plant breeding program;
KW inheritance; desirable characteristic; growth; development;
KW disease resistance; environmental adaptability; quality; yield;
KW multigene trait; plant; gene; ss.
XX
OS Zea mays.
XX
PN US6476212-B1.
XX
PD 05-NOV-2002.
XX
PF 14-MAY-1999; 99US-00313294.
XX
PR 26-MAY-1998; 98US-0086722P.
XX
PA (INCY-) INCYTE GENOMICS INC.
XX
PI laljudi RV, Ito LY, Sherman BK;
XX
DR WPI; 2003-208840/20.
XX
PT Novel purified corn-ear derived polynucleotide useful as hybridization
PT probe for detecting polynucleotide in sample, and for identifying,
PT evaluating, and altering desired characteristics associated with growth,
PT development.
XX
PS Example; SEQ ID NO 121; 390bp; English.
XX
CC The present invention relates to the isolation of corn ear-derived
CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
CC and SATMON023. Some of the cdps uniquely identify structural, functional,
CC and regulatory genes of corn ear. The polynucleotides sequences are
CC useful for detecting cdps in a sample, for producing a corn ear-specific
CC profile of gene transcription, for detecting altered gene expression in
CC inbred or hybrid plants, and for screening several molecules for specific
CC binding to the polynucleotide. The cdps are useful to identify, isolate,
CC or extend identical or related corn-ear nucleic acid sequences from DNA
CC libraries, and in nucleic acid amplification or hybridisation techniques
CC to follow the expression of desirable traits through plant breeding
CC programs. Preferably, the cdps are used to identify, evaluate, alter, or
CC follow the inheritance of desired characteristics associated with growth
CC and development, disease resistance, environmental adaptability, quality,
CC and yield of corn. The cdps are also useful as molecular markers for
CC studying inheritance and multigene traits in a plant breeding program.
CC The cdps are useful for producing purified corn-ear polypeptides by
CC recombinant techniques. They are also useful in diagnostic assays to
CC detect or confirm conditions or diseases associated with abnormal levels
CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
CC polynucleotides (cdps) of the invention. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from the USPTO web site at
CC seqdata.uspto.gov/psipsoIDEntry.html
XX
SQ Sequence 272 BP; 79 A; 52 C; 68 G; 67 T; 0 U; 6 Other;

```

Query Match 4.0%; Score 75; DB 7; Length 272;
 Best Local Similarity 57.6%; Pred. No. 8.7e-11;
 Matches 148; Conservative 0; Mismatches 106; Indels 3; Gaps 1;

QY 1009 TATGAGGCTTTAGAGCAGCCATTCTATGCTCCACAGGGCCATGTGGAGTGT 1068
 DB 2 TATGAAGCAAGTAAACACGCTCATGCTTANMGCTCTGGGACCGGTGTATGACG 61
 QY 1069 CTGATGTTGAGAGCAGTGCCATGCTATTTGAGGCGGAGCCCTCCACGGGAGTTA 1128
 DB 62 TTAATATTTGAAGCTCAGCTGTGGGTACATGAAAGCTGAAGCTTGCTGATAAAGCTTT 121
 QY 1129 GCTGAGATGGGTTAGATGAATTCCTG3---GCTCAGAGGCGAGTATGTTTCTGGA 1185
 DB 122 GTTAAACAGGTACAGAGATTCATGSCACCTACGANGTTGATTTGGCTGCT 181
 QY 1186 GGTGTTGCCCACTGATGCTTCTTGCAAGCAAGATGTGACATATTCATCA 1245
 DB 182 GGGAAAGGCGCTATGCTTCTTATGCAAAAGAGATATGACGATTTAAACAAG 241
 QY 1246 CACTCTCAGGCAAAAC 1262
 DB 242 CAGCCATGATTAAGC 258

RESULT 4
 ABX82003
 ID ABX82003 standard; cDNA; 274 BP.
 XX ABX82003;
 XX 24-APR-2003 (first entry)
 XX
 DE Corn ear-derived polynucleotide (cpd) #463.
 XX
 XX Corn ear-derived polynucleotide; cpd; cDNA library; SATMON022; SATMON023;
 KW structural gene; functional gene; regulatory gene;
 KW corn ear-specific profile; gene transcription; gene expression;
 KW hybrid plant; desirable trait expression; plant breeding program;
 KW inheritance; desired characteristics; growth; development;
 KW disease resistance; environmental adaptability; quality; yield;
 KW multigene trait; plant; gene; ss.
 OS
 XX
 XX Zea mays.
 OS
 XX
 XX US6476212-B1.
 XX
 XX 05-NOV-2002.
 XX
 XX 14-MAY-1999; 99US-00313294.
 XX
 XX 26-MAY-1998; 98US-0086722P.
 XX
 XX (INCY-) INCYTE GENOMICS INC.
 XX
 XX Lalgudi RV, Ito LY, Sherman BK;
 XX
 XX WPI; 2003-20840/20.
 XX
 XX Novel purified corn-ear derived polynucleotide useful as hybridization
 PT probe for detecting polynucleotide in sample, and for identifying,
 PT evaluating, and altering desired characteristics associated with growth,
 PT development.
 XX
 XX Example; SEQ ID NO 463; 390pp; English.
 XX
 XX The present invention relates to the isolation of corn ear-derived
 CC polynucleotides (cdps) from the corn (Zea mays) cDNA libraries SATMON022
 CC and SATMON023. Some of the cdps uniquely identify structural, functional,
 CC and regulatory genes of corn ear. The polynucleotides sequences are
 CC useful for detecting cdps in a sample, for producing a corn ear-specific
 CC profile of gene transcription, for detecting altered gene expression in
 CC inbred or hybrid plants, and for screening several molecules for specific

CC binding to the polynucleotide. The cdps are useful to identify, isolate,
 CC or extend identical or related corn-ear nucleic acid sequences from DNA
 CC libraries, and in nucleic acid amplification or hybridization techniques
 CC to follow the expression of desirable traits through plant breeding
 CC programs. Preferably, the cdps are used to identify, evaluate, alter, or
 CC follow the inheritance of desired characteristics associated with growth
 CC and development, disease resistance, environmental adaptability, quality,
 CC and yield of corn. The cdps are also useful as molecular markers for
 CC studying inheritance and multigene traits in a plant breeding program.
 CC The cdps are useful for producing purified corn-ear polypeptides by
 CC recombinant techniques. They are also useful in diagnostic assays to
 CC detect or confirm conditions or diseases associated with abnormal levels
 CC of cdp expression. ABX81541-ABX89140 represent corn ear-derived
 CC polynucleotides (cdps) of the invention. Note: The sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from the USPTO web site at
 CC seqdata.uspto.gov/psipdsIdentify.html
 XX
 XX SQ Sequence 274 BP; 83 A; 49 C; 73 G; 67 T; 0 U; 2 Other;

Query Match 4.0%; Score 74.4; DB 7; Length 274;
 Best Local Similarity 60.5%; Pred. No. 1.3e-10;
 Matches 156; Conservative 0; Mismatches 98; Indels 4; Gaps 2;

QY 1009 TATGAGGCTTTAGAGCAGCCATTCTATGCT-CCACAGGGCCATCGTGGATGAGTGT 1067
 DB 2 TATGAAGCAAGTAAACACGCTCATGCTTANMGCTCTGGGACCGGTGTATGACG 61
 QY 1068 TCTGATGTTGAGAGCAGTGCATGCTATTTGAGGCGGAGCGCTTCCACCGGAGTT 1127
 DB 62 GTTAAATATTTGAAGCTCAGCTGTGGGTACATGAAAGCTGAAGCTTCTCATTAACACTT 121
 QY 1128 AGCTGATGGGGTTAGATGAATTCCTG3---GGTCAAGAGCGAGATATGTTTCTG 1184
 DB 122 TGTATATCAAGGTACAGACAGATTCATGACCTTACGCAAGTTGATTTGCTG 181
 QY 1185 AGGTGTTGCCCACTGATGCTTCTTGCAAGCAAGATCTGACATATTCATCA 1244
 DB 182 TGGGAAAGGCAAGCTATATGTTTCTTACAAAGAGATATGAGGCATTTAAACA 241
 QY 1245 ACACTCTCAGGCAAAAC 1262
 DB 242 GCATTGCCAGGTAAGC 259

RESULT 5
 ABL71922
 ID ABL71922 standard; cDNA; 310 BP.
 XX ABL71922;
 XX 14-MAY-2002 (first entry)
 XX
 XX Corn tassal-derived polynucleotide (cdps) SEQ ID NO:1296.
 DE
 XX
 XX Corn; corn tassal-derived polynucleotide; cdps; hybrid breeding; CDPS;
 KW inheritance; characteristic; growth; development; disease resistance;
 KW environmental adaptability; quality; yield; molecular marker;
 KW multigene trait; plant breeding; corn tassal; gene; ss.
 OS
 XX
 XX Zea mays.
 OS
 XX
 XX US2001051335-A1.
 XX
 XX 13-DEC-2001.
 XX
 XX 16-APR-1999; 99US-00294093.
 XX
 XX 21-APR-1998; 98US-0082567P.
 XX
 XX (IAGC/) LALGUDI R V.
 XX (ITOL/) ITO L Y.
 XX (SHER/) SHERMAN B K.


```

QY 1278 GTTGAATCATACCAAGATGGTTGTTAAAGAGCTGAGGAGATCTCTGAGACAATCA 1337
DB 1860 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACACAGACAGATGACGA 1919
QY 1338 GCAGCTGAACCTATTAAAGCAAGCTCTCAAAACAGAACAGACAGCCAGGTCCTTGA 1397
DB 1920 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACACAGACAGATGACGA 1979
QY 1398 GGAATCTCTGAAATTTATGACGACGAAAGCTGCTGAACCTGACAGAGATATCGGATCG 1457
DB 1980 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACAGACAGATGACGA 2039
QY 1458 GAGACAGAGAACTTAAGATGACATGAACAGAACAGGAAAGATGATGACACAGCAG 1517
DB 2040 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACACAGACAGATGACGA 2099
QY 1518 GTTTTTCATGATTCATCAACAGATCCATGAAAGAGACGCAAGAGAGAAATTT 1577
DB 2100 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACAGACAGATGACGA 2159
QY 1578 CGAGATGTTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1637
DB 2160 GGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2219
QY 1638 TCCCTCTAGCAATGACGATTCGCCAAAGAGCTGAGAAAGTGTCAAAGCTTCAGCGATT 1697
DB 2220 GCAGCAGACGACGACGATGACGACGACGACGACGACGACGACGACGACGACGACG 2279
QY 1698 TCAGAGAAAGAGATGAGAGAGATTTGTGAAAGAGAGAGATCTGATTAAGATCAAG 1757
DB 2280 GCAGCAGAGAGAGACAGACAGAGATTAAGAGAGACGACGACGACGACGATTAAGATCA 2339
QY 1758 GAAGAGATGAGAAAGATGAAGAGAGCATCAAGAGAGATTTTGTGAGAAAGA 1817
DB 2340 GCAGGAGATTAGAGACAGAGAGATTTAGAGAGACGACGACGACGACGAGATTAGAGAGCA 2399
QY 1818 ATTGATGAGGCTTTGGAACAG 1839
DB 2400 GCAGCAGAGATTAGAGAGCAG 2421

RESULT 7
AAF82901
ID AAF82901 standard; DNA; 3489 BP.
XX
AC AAF82901;
XX
DT 11-SEP-2003 (revised)
DT 06-AUG-2003 (revised)
DT 29-JUN-2001 (first entry)
XX
DE Nucleotide sequence of KSHV tethering protein, LANA.
XX
KW Histone H1; tethering protein; LANA; gene therapy; multiple sclerosis;
KW Parkinson's disease; Huntington disease; diabetes; human herpesvirus 8;
KW KSHV; latency-associated nuclear antigen; LANA; ds.
XX
OS Human herpesvirus 8.
XX
FH Key Location/Qualifiers
FT CDS 1..3489
FT /tag= a
XX
PN WO200125484-A2.
XX
PD 12-APR-2001.
XX
PF 29-SEP-2000; 2000MO-US026908.
XX
PR 01-OCT-1999; 99US-00410399.
XX
PA (UNMI ) UNIV MICHIGAN.

```

```

XX
PI Robertson ES, Cotter MA;
XX
DR WPI; 2001-281736/29.
XX
DR P-PSDB; AAF82331.
XX
PT A composition for use in gene therapy comprises an expression vector that
PT includes a nucleic acid sequence encoding a nucleic acid binding protein.
XX
PS Disclosure; Fig 9A; 60pp; English.
XX
XX
CC The invention provides a composition comprising nucleic acid, histone H1
CC protein and expression vector operationally encoding a protein suitable
CC for tethering the nucleic acid to the histone H1 protein, where the
CC tethering protein is LANA. The composition is useful in aiding the
CC retention of the viral DNA in the host cell. The viral vector encodes a
CC protein suitable for tethering DNA to Histone H1. Methods for screening
CC for compounds which are agonistic or antagonistic for the tethering of
CC viral proteins to histone H1 and DNA binding sites are useful for
CC developing the method of viral transfer. The composition has applications
CC to gene therapy, including the treatment of multiple sclerosis,
CC Parkinson's disease, Huntington disease and diabetes. The present
CC sequence represents the nucleotide sequence of the Kaposi's sarcoma
CC associated herpesvirus (human herpesvirus 8) latency-associated nuclear
CC antigen (LANA), which acts as a tethering protein. (Updated on 06-AUG-
CC 2003 to correct OS field.) (Updated on 11-SEP-2003 to standardise OS
CC field)
XX
SQ Sequence 3489 BP; 1053 A; 862 C; 1137 G; 437 T; 0 U; 0 Other;
XX
Query Match 3.1%; Score 58.8; DB 4; Length 3489;
Best Local Similarity 43.4%; Pred. No. 1.4e-05;
Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;
QY 1218 GAAGCAGATCTGACATATTCATCAACATCTTCAAGGCAAAACAGGCTGAATTCGA 1277
DB 1800 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACGACGACGACGATGACGA 1859
QY 1278 GTTGAATCATACCAAGATGTTGTTAAAGAGCTGAGAGCATCTGAGACAAATCA 1337
DB 1860 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACGACGACGACGATGACGA 1919
QY 1338 GCAGCTGAACCTATTAAAGCAAGCTCTCAAAACAGAACAGACGACCAAGTCTTGA 1397
DB 1920 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACACAGACAGATGACGA 1979
QY 1398 GGAATCTCTGAAATTTATGACGACGAAAGCTGCTGAACCTGACAGAGATATCGGATCG 1457
DB 1980 GCAGCAGATGACACAGACAGATGACGACGACGACGACGATGACAGACAGATGACGA 2039
QY 1458 GAGACAGAGAACTTAAGATGACATGAACAGAACAGGAAAGATGATGACACAGCAG 1517
DB 2040 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACACAGACAGATGACGA 2099
QY 1518 GTTTTTCATGATTCATCAACAGATCCATGAAAGAGACGCAAGAGAGAAATTT 1577
DB 2100 GCAGCAGATGACACAGACAGATGACGACGACGACGATGACACAGACAGATGACGA 2159
QY 1578 CGAGATGTTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1637
DB 2160 GGATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 2219
QY 1638 TCCCTCTAGCAATGACGATTCGCCAAAGAGCTGAGAAAGTGTCAAAGCTTCAGCGATT 1697
DB 2220 GCAGCAGACGACGACGATGACGACGACGACGACGACGACGACGACGACGACGACG 2279
QY 1698 TCAGAGAAAGAGATGAGAGAGATTTGTGAAAGAGAGAGATCTGATTAAGATCAAG 1757
DB 2280 GCAGCAGAGAGAGACAGACAGAGATTAAGAGAGACGACGACGACGACGATTAAGATCA 2339
QY 1758 GAAGAGATGAGAAAGATGAAGAGAGCATCAAGAGAGATTTTGTGAGAAAGA 1817
DB 2340 GCAGGAGATTAGAGACAGAGAGATTTAGAGAGACGACGACGACGACGAGATTAGAGAGCA 2399

```

[illegible]

QY	1278	GTTCGAATCTATCCACCAAGATGGTGTGAAAGGAGCTGAGCAGATCTTGAGGCAATCA	1337
Db	1860	GCACGACGGATGACGACGACGAGGATGACGACGACGACGATGACGACGACGACGATGACGA	1919
QY	1338	GCAGCTGAATCTTATTAAGAACAGCTCTTCMAAACMAACAGCAGCGCAAGGTGCTTGA	1397
Db	1920	GCACGACGATGACGACGACGAGGATGACGACGACGACGAGGATGACGACGACGAGATGACGA	1979
QY	1398	GGAACTCTGGAATTAATGAGCCGAAAGCTGCGTAGAACTGCAAGGATTAATCGGATCGT	1457
Db	1980	GCACGACGATGACGACGACGACGATGACGACGACGACGACGATGACGACGACGAGATGACGA	2039
QY	1458	GAGACAGAGAACTTAAGATGACGATGAACAGAAACAGGAGAGAGATGATCACAACGACAG	1517
Db	2040	GCACGACGATGACGACGACGACGACGATGACGACGACGACGACGATGACGACGACGAGATGACGA	2099
QY	1518	GTTTTTCATGATTCATCAATCAAAACAGATCCATGAAGAAGAGACGCAAGAGAGAAATTT	1577
Db	2100	GGAGCAGACAGGATGACGACGACGACGACGATGACGACGACGACGACGAGATGACGACGACA	2159
QY	1578	CGAGATGTTGCACGACGACGAAACGTGCCAAGGTTGTTGGCCAGCAGGACGACGAACTTAA	1637
Db	2160	GGATGACACGACGACGACGACGATGACGACGACGACGACGACGATGACGACGACGACGATGA	2219
QY	1638	TCCTCTTAGCAATCAGCATTCGACCAAGAGAGAGCTGAGAAAGTGTCAAGCTTCATCGAGTT	1697
Db	2220	GCACGACGACGACGACGATGACGACGACGACGACGAGATGAACAGGACGACGACGACGACGAGA	2279
QY	1698	TCAGAGAAAGACATGAGAGAGGTTGTGGAAGAAGAGGACATGCTGATTAAGAATCAAGA	1757
Db	2280	GCACGA	2339
QY	1758	GAAAGAAATGGAAGACATGAAGAAGAGGACATCAGAGGAGATATTGATCTGAGGAAAGA	1817
Db	2340	GCAGAGATTGAGAGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGA	2399
QY	1818	ATTGATGAGGCTTTGGACACAG	1839
Db	2400	GGAGCAGGAGTTAGAGGACAG	2421
RESULT 9			
AAV73805/C			
ID	AAV73805	standard; DNA; 32207 BP.	
AC	AAV73805;		
XX			
DT	17-OCT-2003	(revised)	
DT	25-FEB-1999	(first entry)	
XX			
XX			
DE	KSHV LUR DNA (nucleotides 105,301-137,507).		
XX			
KM	Kaposi's sarcoma; acquired immune deficiency syndrome; AIDS; DHFR; Bcl-2;		
KM	dihydrofolate reductase; LUR; long unique region; vaccine; prophylaxis;		
KW	diagnosis; treatment; HHV8; capsid protein IV; tegument protein IV;		
KW	glycoprotein; Kaposin; cyclin D; immediate early protein; IRF; OX-2;		
KM	v-ach; G-protein coupled receptor; FGARAT; ds.		
XX			
OS	Human herpesvirus 8.		
XX			
PN	US5849564-A.		
XX			
PD	15-DEC-1998.		
XX			
PF	29-NOV-1996; 96US-00770379.		
XX			
PR	29-NOV-1996; 96US-00770379.		
XX			
PA	(UYCO) UNIV COLUMBIA NEW YORK.		
PI	Bohenzky RA, Moore PS, Russo JJ, Chang Y, Edelman IS;		
XX			

DR WPI; 1999-069741/06.
 XX Kaposi's sarcoma-associated herpes virus nucleic acid - encodes
 PT di:hydrofolate reductase and is useful for treatment, prophylaxis or
 PT diagnosis of Kaposi's sarcoma.
 XX
 XX
 PS Disclosure; Col 155-182; 109pp; English.
 CC This sequence is a fragment of the Kaposi's sarcoma-associated
 CC herpesvirus (KSHV) LTR (long unique region). This fragment contains
 CC coding regions for ORF65 which encodes capsid protein IV, ORF66, ORF67
 CC which encodes tegument protein IV, ORF68 which encodes a glycoprotein,
 CC ORF69, K12 which encodes Kaposin, K13, ORF72 which encodes cyclin D,
 CC ORF73 which encodes immediate early protein (IEP), K14 which encodes OX-2
 CC (v-adh), ORF74 which encodes G-protein coupled receptor, ORF75 which
 CC encodes tegument protein/FGARAT, K15. KSHV is a new human Herpesvirus
 CC (HHV8) believed to cause Kaposi's sarcoma (KS) which is the most common
 CC form of neoplasm occurring in persons with acquired immune deficiency
 CC syndrome (AIDS). The DHFR protein is useful for vaccination, prophylaxis,
 CC diagnosis and treatment of a subject with Kaposi's sarcoma and for
 CC detecting expression of a DNA virus associated with Kaposi's sarcoma in a
 CC cell. (Updated on 17-OCT-2003 to standardise OS field)
 XX
 XX
 SQ Sequence 32207 BP; 7229 A; 9156 C; 8713 G; 7109 T; 0 U; 0 Other;
 Query Match 3.1%; Score 58.8; DB 2; Length 32207;
 Best Local Similarity 43.4%; Pred. No. 4.5e-05;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;
 QY 1218 GAAGCAAGATCTGGACATATTCATCAACCTCTCAAGGCAAAACAAGCTGAATTCGA 1277
 DB 20197 GCAGCAGATGACAGACAGAGATGACGACGAGATGACAGACAGAGATGACGA 20138
 QY 1278 GTTGAATATCATCCAGAGATGTTGTAAAGAGCTGAGGCAATCTCTGAGACATCA 1337
 DB 20137 GCAGCAGATGACAGACAGAGATGACGACGAGATGACAGACAGAGATGACGA 20078
 QY 1338 GCAGTGAATCTTAAAGAACAGCTCTCAAAACGAACACAGACAGCCAGGCTTGA 1397
 DB 20077 GCAGCAGATGACAGACAGAGATGACGACGAGATGACAGACAGAGATGACGA 20018
 QY 1398 GGAATCTCTGAAATATATGACGAGAAAGCTGCTAGAACTGACAGAGATATCGATCGT 1457
 DB 20017 GCAGCAGATGACAGACAGAGATGACGACGAGATGACAGACAGAGATGACGA 19958
 QY 1458 GACACGAGAACTTAAGATGACATGAAACGAACAGGAAAGATGATGACACACAG 1517
 DB 19957 GCAGCAGATGACAGACAGAGATGACGACGAGATGACAGACAGAGATGACGA 19898
 QY 1518 GTTTTCATGATTCATCAATCAACAGATCCATGAAAGAGACGCAAGGAGGAAATT 1577
 DB 19897 GCAGCAGAGATGACAGACAGAGATGACGAGATGACGACAGAGATGACAGACGA 19838
 QY 1578 CGAGATGTTGACAGACAGAGAAAGTGCAGAGTTGTTGGCCAGACAGACAGAACTTAA 1637
 DB 19837 GATGACGACAGACAGACAGATGACGACGAGATGACGAGATGACAGACAGATGA 19778
 QY 1638 TCCCTTAGCATGACGATTTGCCAAAGAGAGCTGAGAAAGTCTCAAGCTTCAATCA 1697
 DB 19777 GCAGCAGACAGAGATGACAGACAGAGATGACGAGATGACGAGAGAGAGACGA 19718
 QY 1698 TCAGAGAAAGAGATGAGAGAGTTTGTGAAGAGAGAGAGATGCTGATAAAGATCAAG 1757
 DB 19717 GCAGCAGAGAGACAGAGAGATGACGAGATGACGAGAGATGACAGATGACAG 19658
 QY 1758 GAAGAAATGAAAGACATGAAGAGAGCATCAAGAGATATTTGATCTGGAGAAAG 1817
 DB 19657 GCAGAGATTAGAGACAGAGAGATGACGAGATGACGAGAGATGACAGAGATGA 19598
 QY 1818 ATTGATGAGGCTTTGGAACG 1839
 DB 19597 GCAGCAGAGATTAGAGAGACG 19576

RESULT 10
 AAV19941/C
 ID AAV19941 standard; DNA; 137507 BP.
 XX
 XX
 AC AAV19941;
 XX
 DT 27-AUG-2003 (revised)
 DT 03-AUG-1998 (first entry)
 XX
 DE KSHV long unique coding region and terminal repeat.
 XX
 XX KSHV; HHV8; human herpes virus 8; macrophage inflammatory protein II;
 XX interleukin-6; IL-6; interferon regulatory factor; rheumatoid arthritis;
 XX complement-binding protein; glycoprotein; capsid protein IV; infection;
 XX immediate early protein; Kaposi's sarcoma; protective vaccine; lymphoma;
 XX lymphoproliferative disease; leukaemia; splenomegaly; mycosis fungoides;
 XX HIV immune status; anti-inflammatory agent; therapy; ds.
 XX
 OS Human herpesvirus 8.
 XX
 FH Location/Qualifiers
 FT 1142..2794
 FT /tag= a
 FT /product= "complement-binding protein"
 FT CDS
 FT /tag= b
 FT /product= "glycoprotein B"
 FT CDS
 FT /tag= c
 FT /product= "interleukin 6"
 FT CDS
 FT /tag= d
 FT /product= "macrophage inflammatory protein II"
 FT CDS
 FT /tag= e
 FT /product= "interferon regulatory factor 1"
 FT CDS
 FT /tag= f
 FT /product= "protein T1.1"
 FT CDS
 FT /tag= g
 FT /product= "glycoprotein M"
 FT CDS
 FT /tag= h
 FT /product= "glycoprotein L"
 FT CDS
 FT /tag= i
 FT /product= "interferon regulatory factor 2"
 FT CDS
 FT /tag= j
 FT /product= "interferon regulatory factor 3"
 FT CDS
 FT /tag= k
 FT /product= "glycoprotein X"
 FT CDS
 FT /tag= l
 FT /product= "interferon regulatory factor 4"
 FT CDS
 FT /tag= m
 FT /product= "capsid protein IV"
 FT CDS
 FT /tag= n
 FT /product= "immediate early protein"
 XX
 PN W09804576-A1.
 XX
 PD 05-FEB-1998.
 XX
 PF 22-JUL-1997;
 XX 97WO-US013346.
 XX 25-JUL-1996;
 PR 96US-00686243.
 XX 25-JUL-1996;
 PR 96US-00686349.

PR 25-JUL-1996; 96US-00686350.
 PR 25-JUL-1996; 96US-00687253.
 PR 05-SEP-1996; 96US-00708678.
 PR 10-OCT-1996; 96US-00728323.
 PR 13-NOV-1996; 96US-00747887.
 PR 13-NOV-1996; 96US-00748640.
 PR 29-NOV-1996; 96US-00757669.
 XX (UYCO) UNIV COLUMBIA NEW YORK.
 XX PA
 XX PI Chang Y, Bohenzky RA, Russe UT, Edelman IS, Moore PS;
 XX DR WPI, 1998-130615/12.
 XX
 PT New nucleic acid encoding Kaposi's sarcoma associated herpes virus
 PT proteins - useful for, e.g. detecting levels of HHV8 In, and preparation
 PT of vaccines for treatment of, HIV patients.
 PS
 PS XX Example 2; Page 135-203; 230pp; English.
 CC This sequence represents the long unique region and terminal repeat of
 CC the Kaposi's sarcoma-associated herpes virus (KSHV). KSHV is also known
 CC as human herpes virus 8 (HHV8). This sequence contains the DNAs of the
 CC invention which encode KSHV polypeptides selected from: (a) viral
 CC macrophage inflammatory protein (MIP) II; (b) viral interleukin-6 (IL-6);
 CC (c) viral IRF 1; (d) complement-binding protein; glycoproteins B, M or L;
 CC (d) capsid protein IV encoded by ORF65; and (e) immediate early protein
 CC encoded by ORF73. Labelled probes for the nucleic acid, proteins encoded
 CC by it, and antibodies (Ab) specific for the proteins are useful for
 CC detecting HHV8, specifically for diagnosis of Kaposi's sarcoma, in body
 CC fluids or tissue samples. HHV8 infections can be treated with antisease
 CC or triplex forming molecules or agents that bind specifically to the
 CC protein. Ab may be used for prophylaxis or treatment of HHV8 infection,
 CC while the protein can be used in protective vaccines. Ab may also be used
 CC to differentiate between lymphomas, and HHV8 may be implicated in many
 CC other lymphoproliferative diseases such as lymphomas, leukaemia,
 CC splenomegaly and mycosis fungoides. Cells and animals containing the
 CC nucleic acid are useful for drug screening. HHV8-derived peptides can be
 CC used as targets for antiviral drugs, e.g. dihydrofolate reductase gene
 CC can be inhibited with methotrexate. These can also be used to determine
 CC the immune status of a patient infected with HIV. HHV8 derived protein
 CC viral MIP III may be used as an anti-inflammatory agent for, e.g.,
 CC treating rheumatoid arthritis. This sequence is stated as containing 81
 CC open reading frames. (Updated on 27-AUG-2003 to correct OS field.)
 XX
 XX Sequence 137507 BP; 32579 A; 37795 C; 35758 G; 31375 T; 0 U; 0 Other;
 SQ
 Query Match 3.1%; Score 58.8; DB 2; Length 137507;
 Best Local Similarity 43.4%; Pred. No. 9.7e-05;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;
 QY 1218 GAAGCAAGATCTGGACATATTCAATCAACTCTCAAGGAAACAAAGGCTGAATTCGA 1277
 DB 125497 GCAGCGAGATGACGACGACGATGACGCGAGATGACGACGATGATGACGA 125438
 QY 1278 GTTGAATCATATCAAGAGATGTTTAAAGAGCTGAGGAGATCTCTGAGACATCA 1337
 DB 125437 GCAGCGAGATGACGACGACGATGACGCGAGATGACGACGATGATGACGA 125378
 QY 1338 GCAGCTGAACCTATTAAAGCAAGCTCTCAAAACGAACAGCAGCCAGGCTGTTGA 1397
 DB 125377 GCAGCGAGATGACGACGACGATGACGCGAGATGACGACGATGATGACGA 125318
 QY 1398 GGAATCTCGGAATTAATGAGCGAAGAGCTGCTAAGCTGAGAGATTAATGATTCGT 1457
 DB 125317 GCAGCGAGATGACGACGACGATGACGCGAGATGACGACGATGATGACGA 125258
 QY 1458 GAGACAGAGAATTAAGATGACGATGAACGAGACGAGAGATGATGACGACAGAG 1517
 DB 125257 GCAGCGAGATGACGACGATGACGCGAGATGACGCGAGATGACGACGATGACGA 125198
 QY 1518 GTTTTCATGATTCATCAACAGATCCATGAAAGAGAGCGAAAGAGAGGAATTT 1577

DB 125197 GCAGCGACGAGATGACGACGACGACGATGACGAGACGACGATGACGACGACGA 125138
 QY 1578 CGAGATGTTGACGACGACGAGAACTGCAAGGTTGTCGCCGACGACGACAGAACTTAA 1637
 DB 125137 GCATGACGACGACGACGACGATGACGACGACGACGACGATGACGACGACGATCA 125078
 QY 1638 TCCCTCTAGCAATGACGATTCGCGAAAGAGAGCTGAGGAAGTGTCAAGCTTCATGAGTT 1697
 DB 125077 GCAGCGACGACGACGATGACGACGACGACGATGAAACGAGGACGACGACGACGAG 125018
 QY 1698 TCAAGAGAAAGAGATGAGAGATTTGTGGAAGAGGAGATGCTGATTAAGATCAAGA 1757
 DB 125017 GCAGCGAGAGGACGACGACGACGACGATTAAGAGAGAGAGGATTAAGAGATCA 124958
 QY 1758 GAAGAAGATGAAGAATGAAAGAGAGGATGACGACGAGATATTGATCTGGAAGAA 1817
 DB 124957 GCAGGAGTTAGAGAGACGACGACGACGATTAAGAGAGAGAGGATTAAGAGAGCA 124898
 QY 1818 ATTTGATGAGGCTTTGGAACAG 1839
 DB 124897 GCAGCGAGGATTAAGAGAGCAG 124876
 RESULT 11
 ID AB199537 standard; cDNA; 1080 BP.
 XX AB199537;
 AC AB199537;
 XX 07-MAR-2002 (first entry)
 DE Mouse ischaemic condition related cDNA sequence SEQ ID NO:537.
 XX
 KW Mouse; ischaemia; compressive ischaemia; occlusive ischaemia;
 KW vasospastic ischaemia; ischaemic condition; ischaemic disease; ss.
 XX
 OS Mus musculus.
 XX
 PN WO200188188-A2.
 XX
 PD 22-NOV-2001.
 XX
 PF 18-MAY-2001; 2001WO-0P004192.
 XX
 PR 18-MAY-2000; 2000JP-00145977.
 XX
 PA (UNIV-) UNIV NIHON SCHOOL JURIDICAL PERSON.
 XX
 PI Ishikawa K, Asai S, Takahashi Y, Nagata T, Ishii Y;
 XX
 P-PSDB; ABB57221.
 XX
 PT Examining the ischaemic condition (e.g. occlusive ischaemia) by measuring
 PT expression levels of particular genes defined in the specification or by
 PT determining the expression profile of a gene group comprising these
 PT genes.
 XX
 PS Claim 2; Page 1472-1473; 2690pp; English.
 XX
 CC The present invention describes a method for examining ischaemic
 CC conditions, comprising measuring the expression levels of particular
 CC genes (I) in a test sample or determining the expression profile of a
 CC gene group in the sample comprising genes selected from (I). The method
 CC is useful for examining the ischaemic condition (e.g. compressive
 CC ischaemia, occlusive ischaemia or vasospastic ischaemia) by measuring
 CC expression levels of particular genes (AB199202 to AB199912, encoding the
 CC protein sequences in ABB57020 to ABB57374) or by determining the
 CC expression profile of a gene group comprising these genes. The expression
 CC levels or expression profiles produced by these genes are used as an
 CC indicator when screening for ischaemic condition-improving drugs or
 CC therapeutics for ischaemic diseases. AB199913 and AB199914 represent PCR
 CC primers for a mouse ischaemic condition related sequence, which are used

CC in the exemplification of the present invention
XX Sequence 1080 BP; 370 A; 191 C; 390 G; 129 T; 0 U; 0 Other;

Query Match 3.0%; Score 56.2; DB 6; Length 1080;
Best Local Similarity 48.0%; Pred. No. 4.1e-05;
Matches 160; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 1466 GAACCTAAGATGACGATGAAACAGACAGGGAAGATGATGACACGACAGGTTTTTCA 1525
DB 701 GAAAG 760
QY 1526 TGGATTCATCAATCAATCATGATCAATGAAAGAGAGAGAGAGAGAGAGATTTTGAATGT 1585
DB 761 AG 820
QY 1586 TGCAGACAGCAG 1645
DB 821 AG 880
QY 1646 GCAATGACGATTCGAG 1705
DB 881 AG 940
QY 1706 AAGAGATGAG 1765
DB 941 AG 1000
QY 1766 TGGAGACATGAAAG 1798
DB 1001 AG 1033

RESULT 12

AA087587
ID AA087587 standard; DNA; 1686 BP.

AC AA087587;

DT 27-AUG-2003 (revised)
DT 19-DEC-1995 (first entry)

DE DNA encoding Leucocytozoan protozoa structural protein epitope.

KM leucocytozoan protozoa; structural protein; epitope; vaccine; fowl;
KW leucocytozoanosis; treatment; ss.

XX Leucocytozoan.

OS JP07089995-A.

PN 04-APR-1995.

XX 10-SEP-1993; 93JP-00226078.

PR 10-SEP-1993; 93JP-00226078.

XX (DOBU-) DOBUTSUO SEIBUTSUGAKUTEKI SEIZAI KYOKAI.
PA (NIRS-) NIRSSEIKEN KK.

DR WPI; 1995-167252/22.
DR P-PSDB; AAR70491.

PT Immune inducing polypeptide against Leucocytozoan protozoa - useful in
PT production of vaccines for treatment of Leucocytozoanosis in fowl.

XX Claim 1; Page 12-14; 20pp; Japanese.

CC AA087587-89 encode polypeptides having a whole or partial epitope of a
CC structural protein of Leucocytozoan protozoa (see AAR70491-93). The
CC polypeptides and DNA encoding them are useful in the production of
CC vaccines for the treatment of Leucocytozoanosis of fowl. (Updated on 27-
CC AUG-2003 to correct OS field.)

XX Sequence 1686 BP; 915 A; 129 C; 382 G; 260 T; 0 U; 0 Other;

Query Match 2.9%; Score 55; DB 2; Length 1686;
Best Local Similarity 44.4%; Pred. No. 0.00011;
Matches 220; Conservative 0; Mismatches 275; Indels 0; Gaps 0;

QY 1305 AAAGAGCTGAGGAGATCTTGAGGACATGACAGCTGAACTTAATTAAGAACAGCT 1364
DB 88 AATGTGAAAG 147
QY 1365 CTCGAAACAG 1424
DB 148 CAAG 207
QY 1425 GCTGCTAG 1484
DB 208 CAAG 267
QY 1485 ACAG 1544
DB 268 GAG 327
QY 1545 CCAAG 1604
DB 328 GAG 387
QY 1605 CAG 1664
DB 388 AATGAG 447
QY 1665 GAG 1724
DB 448 GAG 507
QY 1725 GAG 1784
DB 508 GATGAG 567
QY 1785 GCATCAGAGAGAGAT 1799
DB 568 GCTGAGAGAGAGAGAT 582

RESULT 13

AAT91902
ID AAT91902 standard; DNA; 4000 BP.

AC AAT91902;

DT 07-MAY-1998 (first entry)

DE Mannose-1-phosphate transferase gene MNNA.

KM Mannose-1-phosphate transfer; MNNA gene; enzyme; yeast; regulatory gene;
KW human; high mannose type neutral saccharide chain; ss.

OS Saccharomyces cerevisiae.

XX Key Location/Qualifiers
XX FT CDS 418..3953
XX FT /*tag= a
XX FT /product= "MNNA"

PN JP09266792-A.

XX 14-OCT-1997.

XX 29-MAR-1996; 96JP-00075667.

XX 29-MAR-1996; 96JP-00075667.

PA (AGEN) AGENCY OF IND SCI & TECHNOLOGY.

XX
DR WPI; 1997-553460/51.
DR P-PSDB; AAW30763.
XX
PT Positive regulatory gene of mannose-1-phosphate transfer in yeast -
XX useful for high mannose type neutral saccharide chain production.
XX
PS Claim 2; Page 14-17; 23pp; Japanese.
XX
CC This sequence represents the gene of the invention, designated MN4,
CC encoding a protein which positively regulates mannose-1-phosphate
CC transfer in yeast. The gene is useful for the preparation of human high
CC mannose type neutral saccharide chain
XX
SQ Sequence 4000 BP; 1453 A; 750 C; 770 G; 1027 T; 0 U; 0 Other;
Query Match 2.9%; Score 55; DB 2; Length 4000;
Best Local Similarity 47.9%; Pred. No. 0.00018;
Matches 191; Conservative 0; Mismatches 205; Indels 3; Gaps 1;
QY 1441 GAGGATATTCGATCGTGAAGACAGAGAACTAAGATCAGATGACAGAAACAGGAGAG 1500
Db 3539 GAG 3598
QY 1501 ATGATGACACACACAGAGGTTTTCATGATTCATCAACAGATCCATGAAAGAGAGAC 1560
Db 3599 AAG 3655
QY 1561 GCAAG 1620
Db 3656 GAG 3715
QY 1621 CAGCAGCAG 1680
Db 3716 AAGAT 3775
QY 1681 TCAAGCTTCCTCAGTTTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1740
Db 3776 GAG 3835
QY 1741 CTGATTAAGATCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1800
Db 3836 AAGCAG 3895
QY 1801 TTGATCTGAG 1839
Db 3896 GAAAG 3934
RESULT 14
AAC40071
ID AAC40071 standard; DNA; 1140 BP.
XX
XX AAC40071;
XX
XX 17-OCT-2000 (first entry)
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 26940.
XX
XX Hybridization assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX BP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX
XX 05-MAR-1999; 99US-013180P.

PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126268P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132487P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143342P.
PR 14-JUL-1999; 99US-0143624P.

PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144684P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.
 PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148655P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149375P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.

PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0158993P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160981P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.
 PR 25-OCT-1999; 99US-0161404P.
 PR 25-OCT-1999; 99US-0161405P.
 PR 25-OCT-1999; 99US-0161406P.
 PR 26-OCT-1999; 99US-0161359P.
 PR 26-OCT-1999; 99US-0161360P.
 PR 26-OCT-1999; 99US-0161361P.
 PR 28-OCT-1999; 99US-0161920P.
 PR 28-OCT-1999; 99US-0161922P.
 PR 28-OCT-1999; 99US-0161993P.
 PR 29-OCT-1999; 99US-0162142P.

Query Match 2.8%; Score 53.2; DB 3; Length 1140;

Best Local Similarity 45.2%; Pred. No. 0.0003; Indels 0; Gaps 0;

Matches 196; Conservative 0; Mismatches 238; Indels 0; Gaps 0;

QY 1381 CACGCCAGGTCCTGAGCAATCTCTGAAATTTATGAGGAGAGCTGGGTAGAACTGCA 1440
 DB 473 CACGAGGATGATGTCGATGATGTCACAAAACATAGTGTAGATTTATGATGGA 532
 QY 1441 GAGGATTAATCGATCTGAGACAGAGAACTAAGTCAGCATGAAACAGAGGAAAG 1500
 DB 533 GATGAGAGAGATGAGAAAGAGAGAGAGAGAGAGAGAGATTAAGAAAGATGTAAT 592
 QY 1501 ATGATGCAACAGCAGGTTTTTTCATGATTCATTAACATCCATGAAAGAGAC 1560
 DB 593 AACTCTGAAGATGATGATTAAGAAAGAGAGAGAGAGATGATCAAGAGCATCAT 652
 QY 1561 GCAAAGAGAGAAATTTGAGATGTTGACAGCAGAGAACTGCCAAGGTTGTCAG 1620
 DB 653 GATGATGATGATTTATGATGAGAGAGAGAGAGAGAGAGATTAATGATGATGAG 712
 QY 1621 CAGCAGCAGACATTAATCTCTAGCAATGACGATTCGGAAGAGAGCTGAGAA 1680
 DB 713 AAGAGAGAGAGAGATTAATGATGATGATGATGATGAGAGAGAGAGATTAAT 772
 QY 1681 TCAAGCTTATGAGTTTCAAGAGAGAGAGAGAGAGAGATTTGTGAAGAGAGAG 1740
 DB 773 GATGATGATGATGAGAGAGAGAGAGAGAGAGATTAATGATGATGATGATGAG 832
 QY 1741 CTGATTAAGATCAAGAGAGAGAGATGAGAGCATGAGAGAGAGCATGACAGAGATA 1800
 DB 833 AAGAGAGATTAATGATGATGATGATGATGAGAGAGAGAGAGAGATTAATGAT 892
 QY 1801 TTGATCTGAGAA 1814
 DB 893 GATGATGAGAGAA 906

RESULT 15
 AAX83006/c
 ID AAX83006 standard; DNA; 16442 BP.

AU236368 647 bp mRNA linear EST 01-APR-2002
 LOCUS AU236368 RAF1.4 Arabidopsis thaliana cDNA clone RAF1.4-93-K05 5',
 DEFINITION mRNA sequence.
 ACCESSION AU236368
 VERSION AU236368.1 GI:19875537
 KEYWORDS Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 647)
 Seki, Y., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,
 Oono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y.,
 Arakawa, T., Shibata, K., Shinagawa, A., Muramatsu, M., Hayashizaki, Y.
 and Shinozaki, K.
 Large scale analysis of Arabidopsis full-length cDNA (2002b)
 Unpublished (2002)
 Contact: Motoaki Seki
 Plant Functional Genomics Research Group
 RIKEN Genomic Sciences Center
 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
 Tel: 81-298-36-4359
 Fax: 81-298-36-9060
 Email: mseki@rc.riken.go.jp
 An Arabidopsis full-length cDNA library was constructed essentially
 as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
 and XhoI was ligated to modified Lambda FIC-1 vector (Carninci et
 al., submitted for publication) digested with BamHI and SalI. This
 clone is in a modified pluscript vector. Please visit our web
 site (http://www.gsc.riken.go.jp/e/planet/index_e.html) for further
 details.

FEATURES
 source
 1..647
 /location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clone="RAF1.4-93-K05"
 /issue_type="root"
 /lab_host="DH10B"
 /clone_1lb="RAF1.4"
 /note="Site_1: BamHI; Site_2: SalI"

ORIGIN
 Query Match 28.4%; Score 533; DB 9; Length 647;
 Best Local Similarity 99.5%; Pred. No. 8.1e-118;
 Matches 555; Conservative 0; Mismatches 1; Indels 2; Gaps 2;

1 ATAGGTTCTAGGGCTGTCATGCTTAAGGAAAGAAAGCTTCAAGGCTGTTAGGCTT 60
 Db 92 ATGAGTTCTAGGGCTGTCATGCTTAAGGAAAGAAAGCTTCAAGGCTGTTAGGCTT 151
 Qy 61 GAGGTGAACAGTGGTTCAAGGTTTGCGAGGACGAGTGGCTTCTTCAAGAATGAT 120
 Db 152 GAGGTGAACAGTGGTTCAAGGTTTGCGAGGACGAGTGGCTTCTTCAAGAATGAT 211
 Qy 121 GGAAGAGAGTGGAGTCAATTCAGAGAAACAAGAACAAACCAAGAAACACTTGTGA 180
 Db 212 GGAAGAGAGTGGAGTCAATTCAGAGAAACAAGAACAAACCAAGAAACACTTGTGA 271
 Qy 181 AAAAAGTGGGTTCTCAGAAATTCGATCCCTAGAGCTTGGGTTGGTCAAGCAAGG 240
 Db 272 AAAAAGTGGGTTCTCAGAAATTCGATCCCTAGAGCTTGGGTTGGTCAAGCAAGG 311
 Qy 241 AAGAGTAGCAACGATCTGGAGAGAGAAACAATGATCCGGAGAGGTAAACGCAATGCT 300
 Db 332 AAGAGTAGCAACGATCTGGAGAGAGAAACAATGATCCGGAGAGGTAAACGCAATGCT 391
 Qy 301 CGGGGATTCACACTAATATCTGGTCCGGGACGAGCTTGGACAGAAAGTATGATAAC 360
 Db 392 CGGGGATTCACACTAATATCTGGTCCGGGACGAGCTTGGACAGAAAGTATGATAAC 451

Qy 361 AACTTTGTGGACACCCCACTGATCTGCCTCTTTGAAAGAGATGAAATTGGAC 420
 Db 452 AACTTTGTGGACACCCCACTGATCTGCCTCTTTGAAAGAGATGAAATTGGAC 511
 Qy 421 GCAGAGAGAGGTTTCTGCTCAGACACAGCTGTGCAGAGATTTCTTCAAGTGAATGAT 480
 Db 512 GCAGAGAGAGGTTTCTGCTCAGACACAGCTGTGCAGAGATTTCTTCAAGTGAATGAT 571
 Qy 481 GTGATTAATGCTTCTGAGAGAGAAATGATCCGATGCTTTGGATGATCTGTAGACAG 540
 Db 572 GTGATTAATGCTTCTGAGAGAGAAATGATCCGATGCTTTGGATGATCTGTAGACAG 629
 Qy 541 CTTCGAAAGTATGATAT 558
 Db 630 CTTCGAAAGTATGATAT 647

RESULT 4
 A0964580/c 613 bp DNA linear GSS 28-JAN-2000
 LOCUS LERGX20TF LERG Arabidopsis thaliana genomic clone LERGX20, genomic
 DEFINITION survey sequence.
 ACCESSION A0964580
 VERSION A0964580.1 GI:6792281
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 613)
 Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Uterbach, T.,
 Feldblyum, T., Liang, F., Creasy, T. and Fraser, C.M.
 Genomic survey sequencing of Landsberg erecta ecotype of
 Arabidopsis thaliana and identification of sequence-based
 polymorphisms
 Unpublished (2000)
 Contact: Xiaoying Lin
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: atc@igr.org
 For additional information, see <http://www.igr.org/cdb/at/at.html>
 Seg primer: TP
 Class: Shotgun.

FEATURES
 source
 1..613
 /location/Qualifiers
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /strain="LANDSBERG ERCTA"
 /db_xref="taxon:3702"
 /clone="LERGX20"
 /clone_1lb="LERGX20"
 /note="Organ: Leaf; Vector: pUC19/JK; Total genomic DNA was
 sheared to 0.4-0.7 Kbp before ligation."

ORIGIN
 Query Match 27.9%; Score 523.4; DB 28; Length 613;
 Best Local Similarity 99.8%; Pred. No. 1.7e-115;
 Matches 524; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATAGGTTCTAGGGCTGTCATGCTTAAGGAAAGAAAGCTTCAAGGCTGTTAGGCTT 60
 Db 525 ATGAGTTCTAGGGCTGTCATGCTTAAGGAAAGAAAGCTTCAAGGCTGTTAGGCTT 466
 Qy 61 GAGGTGAACAGTGGTTCAAGGTTTGCGAGGACGAGTGGCTTCTTCAAGAATGAT 120
 Db 465 GAGGTGAACAGTGGTTCAAGGTTTGCGAGGACGAGTGGCTTCTTCAAGAATGAT 406
 Qy 121 GGAAGAGAGTGGAGTCAATTCAGAGAAACAAGAACAAACCAAGAAACACTTGTGA 180
 Db 405 GGAAGAGAGTGGAGTCAATTCAGAGAAACAAGAACAAACCAAGAAACACTTGTGA 346

QY 181 AAAACCTGGTTCCTGAGATTCCTCTCTAGAGCTTGGGCTGTACAGCAAGG 240
 Db 345 AAAACCTGGGTTCTCAGAACTTCTCTAGAGCTTGGGCTGTACAGCAAGG 286
 QY 241 AGAGTAGCAAGCTATCTGGAGAGCAATGATCCGGAGAGGTAAAGCAATGAT 300
 Db 285 AGAGTAGCAAGCTATCTGGAGAGCAATGATCCGGAGAGGTAAAGCAATGAT 226
 QY 301 CGGGGCAATCAAGCTATCTGGTCCGGAGAGCAAGCTTGGAGCAAGAAATGATTAAC 360
 Db 225 CGGGGCAATCAAGCTATCTGGTCCGGAGAGCAAGCTTGGAGCAAGAAATGATTAAC 166
 QY 361 AACTTTGGGCAACCCCACTGTATTCGCCCTCTTTGAAAGAGATGAAATTTGGCAG 420
 Db 165 AACTTTGGGCAACCCCACTGTATTCGCCCTCTTTGAAAGAGATGAAATTTGGCAG 106
 QY 421 GGAAGAGAGGTTCTGCTGAGCAACAGCTGTGAGAGATTTCTGACGTGAGATGAT 480
 Db 105 GGAAGAGAGGTTCTGCTGAGCAACAGCTGTGAGAGATTTCTGACGTGAGATGAT 46
 QY 481 GTGATTAATGCTTCTGAGAGAGATGATTCGATGCTTTGGAT 525
 Db 45 GTGATTAATGCTTCTGAGAGAGATGATTCGATGCTTTGGAT 1

RESULT 5
 AV566465/c 619 bp mRNA linear EST 07-SEP-2000
 LOCUS AV566465 Arabidopsis thaliana green siliques Columbia Arabidopsis
 DEFINITION thaliana cDNA clone SQ244b06f 3', mRNA sequence.
 ACCESSION AV566465
 VERSION AV566465.1 GI:8737917
 SOURCE EST.
 ORGANISM Arabidopsis thaliana (thale cress)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 619)
 A large scale analysis of cDNA in Arabidopsis thaliana: Generation
 of 12,028 non-redundant expressed sequence tags from normalized and
 size-selected cDNA libraries
 DNA Res. 7 (3), 175-180 (2000)
 JOURNAL MEDLINE
 PUBMED 10907847
 COMMENT Contact: Erika Asamizu
 The First Laboratory for Plant Gene Research
 Kazusa DNA Research Institute
 Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
 Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
 Location/Qualifiers
 1..619
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /strain="Columbia"
 /db_xref="taxon:3702"
 /clone="SQ244b06f"
 /tissue_type="green siliques"
 /clone_1fb="Arabidopsis thaliana green siliques Columbia"
 /note="Vector: pBluescriptII SK-, Site_1: EcoRI, Site_2:
 XhoI"

FEATURES
 source
 ORIGIN
 Query Match 27.3%; Score 513; DB 9; Length 619;
 Best Local Similarity 100.0%; Pred. No. 5,5e-113;
 Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 CAAGAGATGTTGTAAGAGAGCTGAGCAATCTCTGAGCAATCGACGCTGAATAC 1350
 Db 619 CAAGAGATGTTGTAAGAGAGCTGAGCAATCTCTGAGCAATCGACGCTGAATAC 560

QY 1351 TTTAAGAACAGCTCTCAAAACAGACAGCAGCCCAAGTGTCTTGAGAAATCTCTGAA 1410
 Db 559 TTTAAGAACAGCTCTCAAAACAGACAGCAGCCCAAGTGTCTTGAGAAATCTCTGAA 500
 QY 1411 ATTATGACGAGAAAGCTGCTGTAAGAACTGACAGAGATTAATCGATCTGACAGAACT 1470
 Db 499 ATTATGACGAGAAAGCTGCTGTAAGAACTGACAGAGATTAATCGATCTGACAGAACT 440
 QY 1471 AAGATGACGATGAAACAGAACAGGAGAAAGATGATGACACGACAGGTTTTCATGAT 1530
 Db 439 AAGATGACGATGAAACAGAACAGGAGAAAGATGATGACACGACAGGTTTTCATGAT 380
 QY 1531 TCAATCAACAGATTCATGAAAGAGAGACCCAAAGAGAGAAATTTGAGATGTTGCAG 1590
 Db 379 TCAATCAACAGATTCATGAAAGAGAGACCCAAAGAGAGAAATTTGAGATGTTGCAG 320
 QY 1591 CAGCAGAAACGTCAGAGGTTGTCGACAGACAGCAACAACTTAATCCTCTTGCAAT 1650
 Db 319 CAGCAGAAACGTCAGAGGTTGTCGACAGACAGCAACAACTTAATCCTCTTGCAAT 260
 QY 1651 GACGATTCGCGAAAGAGAGCTGAGAAAGTCTCAAGCTTCATGATTCAGAGAAAGAG 1710
 Db 259 GACGATTCGCGAAAGAGAGCTGAGAAAGTCTCAAGCTTCATGATTCAGAGAAAGAG 200
 QY 1711 ATGAGAGAGTTTGTGAAGAGAGAGAGATGCTGATTAAGATCAAGAGAAAGATGAA 1770
 Db 199 ATGAGAGAGTTTGTGAAGAGAGAGAGATGCTGATTAAGATCAAGAGAAAGATGAA 140
 QY 1771 GACATGAAGAGAGGATCAGAGAGATTTT 1803
 Db 139 GACATGAAGAGAGGATCAGAGAGATTTT 107

RESULT 6
 AQ959659 523 bp DNA linear GSS 28-JAN-2000
 LOCUS AQ959659 LEREM21TR LER Arabidopsis thaliana genomic LEREM21, genomic
 DEFINITION survey sequence.
 ACCESSION AQ959659
 VERSION AQ959659.1 GI:6787360
 KEYWORDS GSS.
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsis.
 1 (bases 1 to 523)
 Buell, C.R., Lin, X., Pai, G., Barnstead, M., Bowman, C., Utterbach, T.,
 Feldblyum, T., Liang, F., Creasy, T., and Fraser, C.M.
 Genomic survey sequencing of Landsberg erecta ecotype of
 Arabidopsis thaliana and identification of sequence-based
 polymorphisms
 Unpublished (2000)
 Contact: Xiaoying Lin
 The Institute for Genomic Research
 9712 Medical Center Dr., Rockville, MD 20850, USA
 Tel: 301 838 0200
 Fax: 301 838 0208
 Email: atc@igf.org
 For additional information, see http://www.tigr.org/cdb/ac/at.html
 Seq primer: TR
 Class: shotgun.
 Location/Qualifiers
 1..523

FEATURES
 source
 ORIGIN
 Query Match 27.3%; Score 513; DB 9; Length 619;
 Best Local Similarity 100.0%; Pred. No. 5,5e-113;
 Matches 513; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1291 CAAGAGATGTTGTAAGAGAGCTGAGCAATCTCTGAGCAATCGACGCTGAATAC 1350
 Db 619 CAAGAGATGTTGTAAGAGAGCTGAGCAATCTCTGAGCAATCGACGCTGAATAC 560

```
FEATURES
source
location/Qualifiers
1..644
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="Jet_neuf"
/db_xref="taxon:3708"
```

RESULT 8	LOCUS	DEFINITION	ACCESSION	VERSION	KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
CD815300		708 bp mRNA linear EST 10-JUL-2003	BN15.025002F020214	BN15	Brassicica napus cDNA clone BN15025002, mRNA sequence.	CD815300						
CD815300			CD815300.1	GI:32497240	EST.	Brassicica napus (rape)						
						Brassicica napus						
						Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Brassica.						
						1 (bases 1 to 708)						
						Genoplante.						
						Genoplante, a major partnership french program in plant genomics						
						Unpublished (2003)						
						Contact: Genoplante						
						Genoplante						

93, rue Henri Rochefort 91025 EVRY CEDEX France
 Tel: 33 1 69 47 54 00
 Fax: 33 1 69 47 54 10
 This sequence has been generated in the framework of the french
 plant genomics programme 'Genoplante' (<http://www.genoplante.com>
 and <http://genoplante-info.infobiogen.fr>).

FEATURES

source

1..708
 /organism="Brassica napus"
 /mol_type="mRNA"
 /cultivar="Jec neuf"
 /db_xref="taxon:3708"
 /clone="BN15025002"
 /tissue_type="seed"
 /clone_lib="BN15"

ORIGIN

Query Match 23.0%; Score 431.2; DB 14; Length 708;
 Best Local Similarity 78.1%; Pred. No. 3,6e-93;
 Matches 575; Conservative 0; Mismatches 118; Indels 43; Gaps 3;

QY 1057 GGGATGAGTGTCTGATGTTTGAAGACATGCACTGGCTATTGAGGCCGAGCCCTC 1116
 Db 1 GGGATTAGTCTTGAATCTTTGAGATACCGCCACTGGCTACTTTGAGCTGACCGCTTA 60
 QY 1117 CAC-CGGAGTTAGCTGAGATGGGCTTGAATTAATTTGCTGGGGTCAGAGCCAGTAT 1175
 Db 61 CACAGGGGACTTGTAGCTCAAAAGGGGTTTGAATAGAGACCTGGAGCCGTTGCCGAGATG 120
 QY 1176 GTTTTCGAGAGTGTGGCCACTGTATGGCTCTCTTGAACAGACGAATCTGACAT 1235
 Db 121 GTTTTCGAGAGTGTGGCCACTGTATGGCTCTCTTGAACAGACGAATCTGACAT 180
 QY 1236 ATTCATCAACACTCTCA--AGGCAAAACAGGCTGAAATTCGAGTTGAATCATACCA 1292
 Db 181 ATTCATCAACACTCTCAAGATGAGGAAACAGGCTGACGATGAGTGAATCTGACCA 240
 QY 1293 AGAGATGTTGTAAAGAGCTGAGGCGATCTCTGAGACAAATCAGAGCTGAATCTT 1352
 Db 241 GAGATAGTGTGAAGAGCTGAGGCGATCTCTGAGACAAATCAGAGCTGAATCTT 300
 QY 1353 TAAGAACAGCTCTCAAAAGAGAACAGACGCGCAAGGCTCTTGAAGATCTCTGAAAT 1412
 Db 301 CAGAACAGCTCTCAAAAGAGAACAGACGCGCAAGGCTCTTGAAGATCTCTGAAAT 360
 QY 1413 TATGAGCGAAGAGCTGCTAGAACTGCAAGAGATATCGATGCTGACAGAACTAA 1472
 Db 361 TTTGAGCGAAGAGCTGCTAGAACTGCAAGAGATATCGATGCTGACAGAACTAA 420
 QY 1473 GATGACAGATGAACAGAACAGAGAGATGATGACACAGACAGGTTTTCATGATTC 1532
 Db 421 GATGACAGATGAACAGAACAGAGAGATGATGACACAGACAGGTTTTCATGATTC 480
 QY 1533 AATCAACAGATCCATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 1592
 Db 481 CATCAACAGATCCATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 540
 QY 1593 GCAAGAACTGCGCAAGGTTGTTGGCCAGCAGCAGACAGACATTAATCCTTACCAATGA 1652
 Db 541 GCAAGAACTGCGCAAGGTTGTTGGCCAGCAGCAGACAGACATTAATCCTTACCAATGA 562
 QY 1653 CGATGCCGAAAGAT 1712
 Db 563 -GATTTCCGAAAGAT 621
 QY 1713 GAGAGAGTTGTGAT 1772
 Db 622 GAAAGAGTTGTGAT 681
 QY 1773 CATGAGAGAGAGAGAT 1788
 Db 682 CTGAAAGAAAAATAT 697

RESULT 9

CF517786

LOCUS

DEFINITION

CF517786 849 bp mRNA linear EST 09-SEP-2003
 CAP0005_IF_E12 Vitis vinifera cv. cabernet sauvignon (clone 8)
 Petiole - CAP Vitis vinifera cDNA clone CAP0005_IF_E12 5', mRNA
 sequence.

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

CF517786.1 GI:34549554
 EST
 Vitis vinifera
 Vitis vinifera
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; Vitaceae; Vitis.
 1 (bases 1 to 849)
 Goes da Silva, F., Iandolo, A., Lim, H., Baek, J., Jones, K. and
 Cook, D.

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
 berries at various developmental stages
 Unpublished (2003)
 Contact: Douglas Cook, PhD
 CAES Genome Facility
 UC Davis, Plant Pathology
 One Shields Ave, Davis, CA 95616, USA
 Tel: 530 754 6561
 Fax: 530 754 6617
 Email: drcoc@ucdavis.edu
 Seq primer: ACGTACCGACATATGCC.
 Location/Qualifiers
 1..849
 /organism="Vitis vinifera"
 /mol_type="mRNA"
 /cultivar="Cabernet Sauvignon"
 /db_xref="taxon:29760"
 /clone="CAP0005_IF_E12"
 /sex="Hermaphrodite"
 /dev_stage="Onset of Veralson (berry softening)"
 /lab_host="DH5alpha"
 /clone_lib="Vitis vinifera cv. cabernet sauvignon (clone
 8) Petiole - CAP"
 /note="Organ: Petiole; Vector: pDON; Site 1: SfiI; Site 2:
 SfiI; CAP is a cDNA library of Vitis vinifera cv.
 'Cabernet Sauvignon' clone 8 petioles. Samples were
 collected on July 10, 2002 from plants on the onset of
 veraison (berry softening). Sampled vines were located at
 the University of California, Davis, Experimental
 Vineyard. cDNAs were made by oligo-dT priming and
 directionally cloned. 5' and 3' adaptors were used in
 cloning as follows:
 5'-AACGAGTGTATCAACGAGAGTGGCCATTAGGCGCG-3' and
 5'-ATTCTGAGAGCCGAGCGGCCACATG-dT(30)NN-3'. Library was
 constructed using the Clontech Creator SMART kit and
 size-selected to contain the 0.5-3 kb size fraction."

FEATURES

source

ORIGIN

Query Match 20.8%; Score 390; DB 14; Length 849;
 Best Local Similarity 68.5%; Pred. No. 3,6e-83;
 Matches 569; Conservative 0; Mismatches 256; Indels 6; Gaps 2;

QY 469 GTGAGAGATGATGTGATATGCTTCTGAGAGAGAGATATCCGATCTTGTGATAT 528
 Db 19 GAGAGATATTAAGAGAGAGATATGATGACATGATGATGATGATGATGATGAT 78
 QY 529 TCTGATACGACCTTGCAAGATATGATGATGATGATGATGATGATGATGATGAT 588
 Db 79 AGTATGATGATGATCTTCACTGATGATGATGATGATGATGATGATGATGATGAT 138
 QY 589 TCAGAAAGCAATATAGTGTTCAAAGATCTTGTGGAGAGCTTGTGATGATGATGAT 648
 Db 139 ACAGGCAAGAT 198
 QY 649 GAGCAGATTAATGAACCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAT 708

```

Db      199 GAGGAATTAACCAACAGCCAGGAGTGGCATTTGCTTCATGCCAAGGTGACCTGGT 258
QY      709 GCCATGATGTGATATA---ACCTGCAACCTCTACTAGCTATCCGAGAGCAAAAAGAGCT 765
Db      259 GCCATTGACATGAGAGGGGTCTCCAACTTATATACACATGCCAAAAGAGATCA 318
QY      766 AGGCGAGTTAAGCTCCATAGAGATTGGCTGAGATTGTTAGAAAAGATCTACAGATAGA 825
Db      319 AAGAGGTCAAATCTTATGAGAGCTTTCGAAACTTCTGATGAGAGATTGTACAGAGG 378
QY      826 GGGCATCTGTCTATCTCTGTGTGTGATTTATGGGCAAGAAAGGTTTGGTGAAGAT 885
Db      379 GGATCTTCTGTGTACAGCTGAGAGGCAATTTGAAAAGTGAAGGTCATATATGAGTCG 438
QY      886 GAAAAGATTATAAATTGCTGTGGCTCCCAATGTCTATCATATGAAATATTAAGCTGAT 945
Db      439 GTTAAAGATCATAGATTGTTTGCCCTCCAAATGGTTATCATTAATGAACACTAGACTTGA 498
QY      946 AAGGAGATTAAGATAGTGGCTCGGATGGGCAACCAAGAGCTGTGAATTAATCTTGAC 1005
Db      499 CAGAGTCTGATAGAGATGATGGATGGGAAATCAAGACTTCTTGATTTATTTACG 558
QY      1006 AAGTATAGGCTCTTAGAGACGCCATTCCTATGTCCACAGGCGCATCTGGAGTAGT 1065
Db      559 TCATATGCTGACATGAAGGCTCGACATTCCTATGTCTCACAAAGGCGCAGAGGAATGAT 618
QY      1066 GTTCTGATGTTTGAAGACAGTGCCTATTTTGAAGCCGCAAGCTCTCCACGGAG 1125
Db      619 GTTTTATTTTATAGAGCTCGGCAATTTGTTACTTGAAGGCTGAAAGCTTCTTAAGCAC 678
QY      1126 TTAGCTGAGATGGGTTAGATAGATTAATGCTTGGGGGTGAGAAAGC---CAGTATGTTTCT 1182
Db      679 TTGCTGAGCAAGGACACAGATAGGAAGCATGGGATGCTGTCTTCTTCTATCT 738
QY      1183 GGAAGTGTTCGCAACTGATAGCTTCTTCTTCAACGAGCAAGATCTGACATATTCAT 1242
Db      739 GGAGATCNGCGCAGCTTATGATCATGGCTATAAAGAAATTTGACATATTCAC 798
QY      1243 CAACACTCTTAAGCAAAACAGGCTGAATTGAGTTGAATCAATCAAC 1293
Db      799 CAACATCTCAAGAAATCAAAATCAAGTTTGAAGTGAATCATATCA 849

```

```

RESULT 10
AV525508 380 bp mRNA linear EST 01-SEP-2000
LOCUS AV525508 Arabidopsis thaliana aboveground organs two to six-week
DEFINITION Old Arabidopsis thaliana cDNA clone APD25402R 5', mRNA sequence.
ACCESSION AV525508
VERSION AV525508.1 GI:8685036
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE 1 (bases 1 to 380)
AUTHORS Asamizu, E., Nakamura, Y., Sato, S. and Tabata, S.
TITLE A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
JOURNAL DNA Res. 7 (3), 175-180 (2000)
MEDLINE 20361093
PUBMED 10907847
COMMENT Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 293-0812, Japan
Email: asamizukazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
source 1..380
/organism="Arabidopsis thaliana"

```

```

/mol_type="mRNA"
/strain="Columbia"
/db_xref="taxon:3702"
/clone="APD25402R"
/issue_type="aboveground organs"
/dev_stage="two to six-week old"
/clone_id="Arabidopsis thaliana aboveground organs two to
six-week old"
/note="Vector: pBluescriptII SK-, Site_1: EcoRI, Site_2:
XhoI"

```

ORIGIN

```

Query Match 20.1%; Score 378.4; DB 9; Length 380;
Best Local Similarity 99.7%; Pred. No. 1.6e-80;
Matches 379; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY      276 ATCCGGAGAGGTACCGCAATGTCGGGCAATTCAGCTAACATATCTGTGGGAGC 335
Db      1 ATCCGGAGAGGTACCGCAATGTCGGGCAATTCAGCTAACATATCTGTGGGAGC 60
QY      336 AGCGTACAGAGATATGATTAACAATTTGGGACACCCACCTATCTGCCCTCC 395
Db      61 AGCGTACAGAGATATGATTAACAATTTGGGACACCCACCTATCTGCCCTCC 120
QY      396 TTGGAAGAGAGATGAAATTTGGCAGCAGAGAGGTTCTGTACACACAGCTGTGA 455
Db      121 TTGGAAGAGAGATGAAATTTGGCAGCAGAGAGGTTCTGTACACACAGCTGTGA 180
QY      456 GGAAGTTCTGACGTGAGAGATGATGATTAATGCTTCTGAGAAAGAAATGATCCGA 515
Db      181 GGAAGTTCTGACGTGAGAGATGATGATTAATGCTTCTGAGAAAGAAATGATCCGA 240
QY      516 TGCTTGAAGATTCGATGACGACCTTCAGATGATTAATGATCGAGATGAGTCA 575
Db      241 TGCTTGAAGATTCGATGACGACCTTCAGATGATTAATGATCGAGATGAGTCA 300
QY      576 AAAGACCATGATCAAGAAAGAGATAGTGTCAAAAAGTTCTTTGGCAGCTTGA 635
Db      301 AAAGACCATGATCAAGAAAGAGATAGTGTCAAAAAGTTCTTTGGCAGCTTGA 360
QY      636 TACCTTGTGATGAGCAGA 655
Db      361 TACCTTGTGATGAGCAGA 380

```

```

RESULT 11
CF517260/c 808 bp mRNA linear EST 09-SEP-2003
LOCUS CF517260 IR_E12 Vitis vinifera cv. cabernet sauvignon (clone 8)
DEFINITION Petiole - CAP Vitis vinifera cDNA clone CAP0005_IR_E12 3', mRNA
sequence.
ACCESSION CF517260
VERSION CF517260.1 GI:34549028
KEYWORDS EST.
SOURCE Vitis vinifera
ORGANISM Vitis vinifera
REFERENCE 1 (bases 1 to 808)
AUTHORS Goes da Silva, F., Iandolo, A., Lam, H., Baek, J., Jones, K. and
Cook, D.
TITLE Expressed sequence tags from Vitis vinifera 'Cabernet sauvignon'
berries at various developmental stages
JOURNAL Unpublished (2003)
COMMENT Contact: Douglas Cook, PhD
CAES Genome Facility
UC Davis, Plant Pathology
One Shields Ave, Davis, CA 95616, USA
Tel: 530 754 6561
Fax: 530 754 6617
Email: drcoc@ucdavis.edu
Seq primer: GCCAACAAGATGCTTAG.

```

FEATURES
source

Location/Qualifiers
1. 808
/organism="Vitis vinifera"
/mol_type="mRNA"
/cultivar="Cabernet Sauvignon"
/db_xref="taxon:29760"
/clone="CAP0005 IR E12"
/sex="Hermaphrodite"
/dev_stage="onset of veraison (berry softening)"
/lab_host="Dh5alpha"
/clone_lib="Vitis vinifera cv. cabernet sauvignon (clone 8) Petiole - CAP"
/note="Organ: Petiole; Vector: pDNR; Site 1: Sfil; Site 2: Sfil; CAP is a cDNA library of Vitis vinifera cv. 'Cabernet Sauvignon' clone 8 petioles. Samples were collected on July 10, 2002 from plants on the onset of veraison (berry softening). Sampled vines were located at the University of California, Davis, Experimental Vineyard. cDNAs were made by oligo-dT priming and directionally cloned. 5' and 3' adaptors were used in cloning as follows:
5'-AACGAGTGTATTAAGCCAGATGGCCATTACGGCCGG-3' and
5'-ATTCTAGAGCGGAGCGCGCCAGATG-AT(30)NN-3'. Library was constructed using the Clontech Creator SMART kit and size-selected to contain the 0.5-3 kb size fraction."

ORIGIN

Query Match 19.4%; Score 364.4; DB 14; Length 808;
Best Local Similarity 68.0%; Pred. No. 5,4e-77;
Matches 538; Conservative 0; Mismatches 247; Indels 6; Gaps 2;

595 AACGAGTGTATTAAGCCAGATGGCCATTACGGCCGG-3' and
808 AAGAGAGCGAGGTGTTAAGTCAATTTTGAAGTATGATGCTTAATCATTTAGAGAA 749
655 ATAAATGACCAAGAGCGAGTGGCATTTGTCAGCTTGTGAGAACGAGCTGTGTCATC 714
748 ATTAAGCAACAGCGAGCGAGTGGCATTTGTCAGCAAGGTGAGCTGTGTCATC 689
715 GATTGTATTA--ACTGCACTCTTACTTACTGATCGAGCAAAAGAGATGAGCA 771
688 GACTGTGACAGGGGTCTCCCAACCTTGATCAACATCCCAAAACAAAGATCAAGAGG 629
772 GTTAAGCTCATGAGAAATGGCTGGAAGTTTGAAGAAAGATCATAGATGAGAGGCA 831
628 GTCAAACTTATCGAGGCTTTCAGAACTTCTGATAGAGATTTGACAGAGGGATCT 569
832 TCTGTCAATCTTGTGTGATGATTTATGCGAGTGAAGGTTTGGGTGAGATGAAG 891
568 TCTGTGTACAGCTGAGAGGCAATTTGGAAGAGTGAAGGTCTAATAGTGGTTAAA 509
892 GATTATGAATTTGCTGCGCTTCCAAATGTCATCATGAAATCTAGATGATGAAGAC 951
508 GATCATATGATTTGTTGGCTTCCAAATGTTATCAATGAACACTAGACTTGAACAGAT 449
952 GATTAAGATAGTGGCTGCGATGGCAACCAAGAGCTGTGAAATCTTCAAGATAT 1011
448 GCTGATAGAGTATGATTTGGATGGGAATCAAGAGCTTCTGATTTATTTCAAGTCAAT 389
1012 GAGGCTTTAGAGACGCCAATCTTATGTTCCACAGGCGCAATGCTGGATGAGTTCTG 1071
388 GCTGATAGAGTATGATTTGGATGGGAATCAAGAGCTTCTGATTTATTTCAAGTCAAT 329
1072 ATGTTTGAAGACAGTGCCTGCTATTTGAGAGCCCAAGCCCTCCACGGAGTATGCT 1131
328 ATTTTGAAGCTCGGCAATTTGTTACTTGAAGCTCAAGCCCTTCAATGAACAATTTGCT 269
1132 GAGATGGGGTTAGTATGATTTGCTGGGGTCAAGAGCG---CAGTATGTTTCTGAGAGT 1188
268 GAGCAAGGACAGATGAGGAAGATGGAGATGCTGTGTTCTTTCTATCTGAGAGGA 209
1189 GTTGGCAACTGATGCTTCTTTCGCAACGAGAGAGATCTGACATATTCATCAACAC 1248

DB 208 AAGCCCGACTTTATGATATACATGGCTATAAGAGAGATTGACATTAATCAACCAACAT 149
QY 1249 TCTAAGGCAAAACAGAGGCTGAATTTGATGTAATCATCAAGAGATGTTGTAAG 1308
DB 148 TCTAAGGAAATTCAAAACCTGAATTTGATGATGATATATCAAGAAATGTTTGT 89
QY 1309 GAGCTGAGGCAATCTCTGAGACATGAGCAAGCTGAATCTTATGAACAAGCTTCA 1368
DB 88 CAATGAGCAAAATAGTAGAGATATATCAACATCAATTTGTTTAAAAACAGAGTTACA 29
QY 1369 AACGACACA 1379
DB 28 AAAAAAAAAA 18

RESULT 12
AI999551/c 548 bp mRNA linear EST 08-SEP-1999
LOCUS 701556368 A. thaliana, Columbia Col-0, rosette-3 Arabidopsis
DEFINITION thaliana cDNA clone 701556368, mRNA sequence.
ACCESSION AI999551 GI:5846456
VERSION AI999551
KEYWORDS EST
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 548)
Chen, J., Momiyama, M., Chan, E., Mooney, M., Carroon, B., Gilliland, D.,
Wang, X., Hillman, J., Guegler, K., Kim, C., Doyle, M., Brzoska, P.,
Gorgone, G., Burns, D., Griffin, J., Mouanoutoua, M., Nguyen, D.,
Tan, R., Rose, M., Warren, B., Ton, B., Kasury, K., Borillo, C.,
Carpio, T., Policky, J., Suzuki, G., Argentine, C., Shah, S.,
Nobrega, A., Murry, L., Turner, C., Krikorian, S., Elder, L. and
Hanson, D.

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Arabidopsis thaliana Gene Expression Microarray
Unpublished (1999)
Contact: David Smoller, Ph.D.
Genome Systems, Inc., a wholly owned subsidiary of Incyte
Pharmaceuticals, Inc.
4633 World Parkway Circle, St. Louis, MO 63134, USA
Tel: 877-577-2733
Fax: 314-427-3324
Email: service@genomesystems.com.

FEATURES
source

Location/Qualifiers
1. 548
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/cultivar="Columbia Col-0"
/db_xref="taxon:3702"
/clone="701556368"
/tissue_type="rosette"
/dev_stage="4 - 7 weeks"
/clone_lib="A. thaliana, Columbia Col-0, rosette-3"
/note="Vector: pSPORT; Site 1: NotI; Site 2: SalI; cDNA library was derived from untreated rosette tissue from Arabidopsis thaliana, Columbia Col-0, at 4 - 7 weeks. Plants were grown in 1.1 l peat moss/vermiculite/perlite soil at 22 deg. C +/- 3 deg. C under constant light, and watered with fertilizer. cDNA synthesis was initiated using a NotI-oligo(dT) primer. Double-stranded cDNA was binned, ligated to SalI adaptors, digested with NotI, size-selected, and cloned into the NotI and SalI sites of the pSPORT vector."

ORIGIN

Query Match 19.2%; Score 359.8; DB 9; Length 548;
Best Local Similarity 97.0%; Pred. No. 6e-76;
Matches 361; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
1507 GCACAGCAGGTTTTCATGATTCATCAACAGATCCATGAAGAGACGCAAG 1566

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 7, 2004, 19:34:49 ; Search time 94.7349 Seconds
(without alignments)
11001.196 Million cell updates/sec

Title: US-10-030-829-2
Perfect score: 1878
Sequence: 1 atgagttctcaggcgtgtcc.....ttccacatgaagatgatga 1878

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database :
1: Issued Patents NA:*
2: /cgn2_6/pcodata/2/ina/5A COMB.seq:*
3: /cgn2_6/pcodata/2/ina/5B COMB.seq:*
4: /cgn2_6/pcodata/2/ina/6A COMB.seq:*
5: /cgn2_6/pcodata/2/ina/6B COMB.seq:*
6: /cgn2_6/pcodata/2/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysts of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	87.8	4.7	7218	1 US-08-232-463-14	Sequence 14, Appl
2	75	4.0	272	4 US-09-313-294A-121	Sequence 121, App
3	74.4	4.0	274	4 US-09-313-294A-463	Sequence 463, App
4	58.8	3.1	3489	2 US-08-728-123A-1	Sequence 1, Appl
5	58.8	3.1	3489	4 US-09-298-568-1	Sequence 1, Appl
6	58.8	3.1	3489	4 US-09-410-399-1	Sequence 1, Appl
C 7	58.8	3.1	32207	2 US-08-770-379-20	Sequence 20, Appl
C 8	58.8	3.1	32207	3 US-08-757-669A-20	Sequence 20, Appl
C 9	58.8	3.1	32207	4 US-09-230-371A-20	Sequence 20, Appl
C 10	53.2	2.8	16442	3 US-08-781-891-208	Sequence 208, App
C 11	53.2	2.8	16442	4 US-09-618-166-208	Sequence 208, App
C 12	49.8	2.7	51259	3 US-08-781-891-209	Sequence 209, App
C 13	49.8	2.7	51259	4 US-09-618-166-209	Sequence 209, App
14	47.4	2.5	1276	3 US-09-411-812A-2	Sequence 2, Appl
15	47.4	2.5	1276	4 US-09-590-113-2	Sequence 2, Appl
16	47.4	2.5	1276	4 US-09-590-113-2	Sequence 2, Appl
17	47.4	2.5	1276	4 US-09-590-113-2	Sequence 2, Appl
18	46.6	2.5	2338	1 US-08-425-069-1	Sequence 268, App
19	46.6	2.5	2338	2 US-08-425-069-1	Sequence 1, Appl
20	45.4	2.4	340	1 US-08-182-175A-104	Sequence 104, App
21	45.4	2.4	340	5 PCT-US92-06412-104	Sequence 104, App
22	43	2.3	1926	4 US-09-249-585A-2	Sequence 2, Appl
23	43	2.3	1926	4 US-09-410-399-3	Sequence 2, Appl
24	43	2.3	2580	3 US-09-050-863-2	Sequence 2, Appl
25	43	2.3	2580	4 US-09-359-081-2	Sequence 2, Appl
C 26	43	2.3	5452	2 US-09-130-114-1	Sequence 1, Appl
C 27	43	2.3	8705	4 US-09-647-344A-14	Sequence 14, Appl

28	43	2.3	9600	3 US-08-910-647-1	Sequence 1, Appl
29	43	2.3	9600	4 US-09-620-925-1	Sequence 1, Appl
30	43	2.3	10596	1 US-07-884-811-15	Sequence 15, Appl
31	43	2.3	10596	1 US-07-885-971-15	Sequence 15, Appl
32	43	2.3	10596	1 US-08-087-783A-15	Sequence 15, Appl
33	43	2.3	10596	1 US-08-194-088B-15	Sequence 15, Appl
34	43	2.3	10596	2 US-08-194-087-15	Sequence 15, Appl
35	43	2.3	10596	5 PCT-US93-04648-15	Sequence 15, Appl
C 36	43	2.3	16080	4 US-09-724-566A-48	Sequence 48, Appl
37	42.2	2.2	1251	4 US-09-828-000-1	Sequence 1, Appl
C 38	42.2	2.2	1270	4 US-08-387-805-1	Sequence 1, Appl
39	42.2	2.2	1958	4 US-09-702-327-3	Sequence 3, Appl
40	41.2	2.2	2277	1 US-08-676-967-2	Sequence 2, Appl
41	41.2	2.2	2277	1 US-08-676-974-2	Sequence 2, Appl
42	41.2	2.2	2277	2 US-09-098-487-2	Sequence 2, Appl
43	41.2	2.2	8920	2 US-08-446-855A-1	Sequence 1, Appl
44	41.2	2.2	8920	3 US-09-150-741-1	Sequence 1, Appl
45	41	2.2	966	2 US-08-766-738-2	Sequence 2, Appl

ALIGNMENTS

RESULT 1
US-08-232-463-14/c
Sequence 14, Application US/08232463
Patent No. 5670367
GENERAL INFORMATION:
APPLICANT: DORNER, F.
APPLICANT: SCHEIFLINGER, F.
APPLICANT: FALKNER, F. G.
TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESSEE: Foley & Lardner
STREET: 1800 Diagonal Road, Suite 500
CITY: Alexandria
STATE: VA
COUNTRY: USA
ZIP: 22313-0299
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/232,463
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/07/935,313
FILING DATE:
APPLICATION NUMBER: EP 91 114 300.6
ATTORNEY/AGENT INFORMATION:
NAME: BENT, Stephen A.
REGISTRATION NUMBER: 29,768
REFERENCE/DOCKET NUMBER: 30472/114 IMMU
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703)836-9300
TELEFAX: (703)683-4109
TELEX: 899149
INFORMATION FOR SEQ. ID NO. 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 7218 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
CLONE: PTZgpc-F1s
US-08-232-463-14
Query Match 4.7%; Score 87.8; DB 1; Length 7218;

Best Local Similarity 4.9%; Pred. No. 2,7e-15;
Matches 20; Conservative 250; Mismatches 137; Indels 0; Gaps 0;

1388 AGGTGCTTGAAGAAATCTGAAATTTATGACGAGAAAGTGGTGAAGTGCAGAGATA 1447
1465 AAGTATTTAAAGATGAGAAATTTGTCACRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1406
1448 ATCGATCTGACGACGAGAACTAATGACATGACGAGAAAGAGATGATG 1507
1405 RRR 1346
1508 CACACGACAGGTTTTCATGATTCATCAATCAACAGATCCATGAAAGAGACGCAAG 1567
1345 RRR 1286
1568 AGGAAATTTGAGATTTGACGACGAGAGAGTCCCAAGTTGTTGCGACGACGAC 1627
1285 RRR 1226
1628 AGAATTAATCCCTTACGAAATGACGATTCGCGAAAGAGCTGAGAGATGTCAGCT 1687
1225 RRR 1166
1688 TCATCGAGTTTCAAGAGAAAGATGAGAGATTTGTCAGAGAGAGAGATGCTATA 1747
1165 RRR 1106
1748 AAGATCAAGAGAGAGATGAGAGATGAGAGAGATGAGAGATGAGAGATGAGAG 1794
1105 RRR 1059

RESULT 2

US-09-313-294A-121
Sequence 121, Application US/09313294A

Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laljudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
NUMBER OF SEQ ID NOS: 1999-05-14
SOFTWARE: PERL Program
SEQ ID NO 121
LENGTH: 272
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700548569H1
NAME/KEY: unsure
LOCATION: 14, 32, 61, 127, 162, 192
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-121

Query Match 4.0%; Score 75; DB 4; Length 272;
Best Local Similarity 57.6%; Pred. No. 2.4e-12;

Matches 148; Conservative 0; Mismatches 106; Indels 3; Gaps 1;

1009 TATGAGCTCTTGAAGACGCGCTTCTTATGTCACAGGGCCATGCGAGATGATG 1068
2 TATGAGCAAGTAAACAGCTCATGCTATNGTCCCTCTGGGACCGGTGATGACGCTN 61
1069 CTGATGTTGAGAGAGTGCCTGCTATTTGAGAGCGGAAAGCCTCCACGGGAGTTA 1128
62 TTAATATTTAAAGCTCAGCTGCGGCTCATGAGAGCTGAGAGCTGATTAACACTTT 121
1129 GCTGAGATGGGTTAGATAGATTTGCTG--GCTCAGAGGCGAGTATTTTCTGGA 1185
122 GTTAANCAAGTACAGACAGAAATTCATGCGACCTAGCGAGGTTGATTTGCTGCT 181

1186 GGTGTCGCCACTGTATGCTTCTTGCACGAGAACAGATGTGACATATTCATCA 1245
182 GGGAAAGGCGCTTATATGTTTCTTAGCAACAAAGAGATGTGACCATTTAAAG 241
1246 CACTTCAAGGCAAAAC 1262
242 CATGCAATGTTAAAGC 258

RESULT 3

US-09-313-294A-463
Sequence 463, Application US/09313294A

Patent No. 6476212
GENERAL INFORMATION:
APPLICANT: Laljudi, Raghunath V.
APPLICANT: Ito, Laura Y.
APPLICANT: Sherman, Bradley K.
TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN EAR
FILE REFERENCE: PL-0017 US
CURRENT APPLICATION NUMBER: US/09/313,294A
NUMBER OF SEQ ID NOS: 1999-05-14
SOFTWARE: PERL Program
SEQ ID NO 463
LENGTH: 274
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
NAME/KEY: misc feature
OTHER INFORMATION: Incyte ID No. 6476212 700549278H1
NAME/KEY: unsure
LOCATION: 15, 53
OTHER INFORMATION: a, t, c, g, or other
US-09-313-294A-463

Query Match 4.0%; Score 74.4; DB 4; Length 274;
Best Local Similarity 60.5%; Pred. No. 3.6e-12;
Matches 156; Conservative 0; Mismatches 98; Indels 4; Gaps 2;

1009 TATGAGCTCTTGAAGACGCGCTTCTTATGTCACAGGGCCATGCGAGATGATG 1067
2 TATGAGCAAGTAAACAGCTCATGCTATNGTCCCTCTGGGACCGGTGATGACGCTN 61
1068 TCTGATTTGAGAGAGTGCCTGCTATTTGAGAGCGGAAAGCCTCCACGGGAGTT 1127
62 GTTAATATTTAAAGCTCAGCTGCGGCTCATGAGAGCTGAGAGCTGATTAACACTTT 121
1128 AGCTGAGTGGGTTAGATAGATTTGCTG--GCTCAGAGGCGAGTATTTTCTG 1184
122 TGTAAATCAAGGTACAGACAGAAATTCATGCGACCTAGCGAGGTTGCTGCTG 181
1185 AGGTGTCGCCACTGTATGCTTCTTGCACGAGAACAGATGTGACATATTCATCA 1244
182 TGGGAAAGGCACTATATGTTTCTTAGCAACAAAGAGATGTGAGGCAATTAACA 241
1245 ACACTTCAAGGCAAAAC 1262
242 GCATTGCGAGTTAAAGC 259

RESULT 4

US-08-728-323A-1
Sequence 1, Application US/08728323A

Patent No. 5948676
GENERAL INFORMATION:
APPLICANT: Chang, Yuan
APPLICANT: Bohenzky, Roy A.
APPLICANT: Russo, James J.
APPLICANT: Edelman, Isidore S.
APPLICANT: Moore, Patrick S.
TITLE OF INVENTION: Immediate Early Protein From Kaposi's
TITLE OF INVENTION: Sarcoma-Associated Herpesvirus, DNA

TITLE OF INVENTION: Encoding Same And Uses Thereof
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Cooper & Dunham LLP
STREET: 1185 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: U.S.A.
ZIP: 10036
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/728,323A
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: White, John P.
REGISTRATION NUMBER: 28,678
REFERENCE/DOCKET NUMBER: 0575/52268/JPM/MS/SKS
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-278-0400
TELEFAX: 212-391-0525
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 3489 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 1..3489
US-08-728-323A-1

Query Match 3.1%; Score 58.8; DB 2; Length 3489;
Best Local Similarity 43.4%; Pred. No. 6.7e-07;
Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

1218 GAAGCAAGATCTGACATATTCATCAATCACTCTCAAGGCAAAACAAGCTGAATTGCA 1277
1800 GCAGCAGAGATGACAGAGAGATGACGACGACGAGATGACAGAGATGACG 1859
1278 GTTGAATCATCAACAGAGATGTTTAAAGAGCTGAGGCAATCTCTGAGGCAATCA 1337
1860 GCAGCAGATGACAGAGATGACGACGAGATGACGAGATGACGAGATGACG 1919
1338 GCAGCTGAATCTTTAAGAAAGCTCTCAAAACGAACAGCAAGCCCAAGGCTTGA 1397
1920 GCAGCAGATGACAGAGATGACGACGAGATGACGAGATGACGAGATGACG 1979
1398 GGAATCTCTGAAATTTATGACGAGAAAGCTGCTAGAACTGACAGGATTAATCGATCGT 1457
1980 GCAGCAGATGACAGAGAGATGACGACGAGATGACGAGATGACGAGATGACG 2039
1458 GAGACAGAGAACTAAGATGACAGTGAACGAACAGGGAAGAGATGATCAACGACG 1517
2040 GCAGCAGATGACAGAGATGACGACGAGATGACGAGATGACGAGATGACG 2099
1518 GTTTTCATGATTCATCAATCAACAGATCCATGAAAGAAAGCAAGCAAGGAGAAATTT 1577
2100 GCAGCAGAGATGACAGAGAGATGACGAGATGACGAGATGACGAGATGACG 2159
1578 CGAGATGTTGACAGCAGAGAGAAAGTGCCTGTTGCTGACAGCAGCAGCAATTTAA 1637
2160 GATGACAGCAGCAGAGATGACGAGATGACGAGATGACGAGATGACGAGATGACG 2219
1638 TCCCTTACGATGACGATGATGCGGAAAGAGCTGAGGAAAGTCAAGCTTCAATCGATT 1697
2220 GCAGCAGCAGCAGAGATGACGAGATGACGAGATGACGAGATGACGAGATGACG 2279

1698 TCAGAGAGAGAGATGAGAGAGATTTGTGAGAGAGAGAGATGCTGATTAAGATCAAGA 1757
2280 GCAGCAGAGAGAGAGAGAGAGATGAGAGAGATGAGAGAGAGAGATGAGAGAGAGAG 2339
1758 GAAGAGATGAGAGAGATGAGAGAGAGATGAGAGAGATGAGAGAGATGAGAGAGAGAG 1817
2340 GAGAGATGAGAGAGAGAGAGAGAGATGAGAGAGAGATGAGAGAGAGAGAGAGAGAGAG 2399
1818 ATTTGATGAGGCTTGAACAG 1839
2400 GCAGCAGAGATGAGAGAGAGAG 2421

RESULT 5
US-09-298-568-1
Sequence 1, Application US/09298568
Patent No. 6322792
GENERAL INFORMATION:
APPLICANT: Kieff, Elliott D.
APPLICANT: Ballester, Mary E.
APPLICANT: Kave, Kenneth M.
TITLE OF INVENTION: RHADINO VIRUS DNA TO MEDIATE EFFICIENT EPISOME PERSISTENCE
FILE REFERENCE: 16412-10001R
CURRENT APPLICATION NUMBER: US/09/298,568
CURRENT FILING DATE: 1999-04-21
EARLIER APPLICATION NUMBER: US 60/109,422
EARLIER FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 3
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 1
LENGTH: 3489
TYPE: DNA
ORGANISM: Kapost's sarcoma-associated herpesvirus
US-09-298-568-1

Query Match 3.1%; Score 58.8; DB 4; Length 3489;
Best Local Similarity 43.4%; Pred. No. 6.7e-07;
Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

1218 GAAGCAAGATCTGACATATTCATCAATCACTCTCAAGGCAAAACAAGCTGAATTGCA 1277
1800 GCAGCAGAGATGACAGAGAGATGACGACGAGATGACGAGATGACGAGATGACG 1859
1278 GTTGAATCATCAACAGAGATGTTTAAAGAGCTGAGGCAATCTCTGAGGCAATCA 1337
1860 GCAGCAGATGACAGAGATGACGACGAGATGACGAGATGACGAGATGACG 1919
1338 GCAGCTGAATCTTTAAGAAAGCTCTCAAAACGAACAGCAAGCCCAAGGCTTGA 1397
1920 GCAGCAGATGACAGAGATGACGACGAGATGACGAGATGACGAGATGACG 1979
1398 GGAATCTCTGAAATTTATGACGAGAAAGCTGCTAGAACTGACAGGATTAATCGATCGT 1457
1980 GCAGCAGATGACAGAGAGATGACGACGAGATGACGAGATGACGAGATGACG 2039
1458 GAGACAGAGAACTAAGATGACAGTGAACGAACAGGGAAGAGATGATCAACGACG 1517
2040 GCAGCAGATGACAGAGATGACGACGAGATGACGAGATGACGAGATGACG 2099
1518 GTTTTCATGATTCATCAATCAACAGATCCATGAAAGAAAGCAAGCAAGGAGAAATTT 1577
2100 GCAGCAGAGATGACAGAGAGATGACGAGATGACGAGATGACGAGATGACG 2159
1578 CGAGATGTTGACAGCAGAGAGAAAGTGCCTGTTGCTGACAGCAGCAGCAATTTAA 1637
2160 GATGACAGCAGCAGAGATGACGAGATGACGAGATGACGAGATGACGAGATGACG 2219
1638 TCCCTTACGATGACGATGATGCGGAAAGAGCTGAGGAAAGTCAAGCTTCAATCGATT 1697
2220 GCAGCAGCAGCAGAGATGACGAGATGACGAGATGACGAGATGACGAGATGACG 2279
1698 TCAGAGAGAGAGATGAGAGAGATTTGTGAGAGAGAGAGATGCTGATTAAGATCAAGA 1757

Db 2280 GCAGCAGAGGAGCAGAGCATGTTAGAGCAGAGCAGAGTTAGAGATAGCA 2339
 QY 1758 GAAGAAGATGGAAGACATGAAGAAGAGCATCGAGAGATATTTGATCTGGAGAA 1817
 Db 2340 GCAGAGATTAGAGAGCAGAGAGGATTTAGAGAGCAGAGAGTTAGAGAGCA 2399
 QY 1818 ATTGATGAGGCTTTGGAAACG 1839
 Db 2400 GGAGCAGAGATTAGAGAGCAG 2421

RESULT 6

US-09-410-399-1
 ; Sequence 1, Application US/09410399
 ; Patent No. 6482587
 ; GENERAL INFORMATION:
 ; APPLICANT: Robertson, Erle S.
 ; APPLICANT: Colter, Murray A.
 ; TITLE OF INVENTION: Methods to inhibit or Enhance the Binding of Viral DNA
 ; FILE REFERENCE: UM-03778
 ; CURRENT APPLICATION NUMBER: US/09/410,399
 ; CURRENT FILING DATE: 1999-10-01
 ; NUMBER OF SEQ ID NOS: 6
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 3489
 ; TYPE: DNA
 ; ORGANISM: Kaposi's sarcoma-associated herpesvirus
 US-09-410-399-1

Query Match 3.1%; Score 58.8; DB 4; Length 3489;

Best Local Similarity 43.4%; Pred. No. 6.7e-07; Mismatches 352; Indels 0; Gaps 0;

Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

QY 1218 GAAGCAAGATCTGACATATTCATCACTCTCAAGGCAAAACAGCTGAAATTCGA 1277
 Db 1800 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 1859
 QY 1278 GTTGAATCATACCAAGATGTTGTTAAAGAGCTGAGGAGATCTTGAGACAATCA 1337
 Db 1860 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 1919
 QY 1338 GCAGCTGAATCTTTAAGAACAAAGCTCTCAAAACAGAAACAGCCAGGCTTGA 1397
 Db 1920 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 1979
 QY 1398 GGAATCTCTGAAATTTATGAGCGAAGAGCTGCGTAGAATGACAGAGATTAATCGATCGT 1457
 Db 1980 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 2039
 QY 1458 GAGACAGAGAACTAAGATGACAGTGAACAGAAACAGGAAAGATGATGACAGCAG 1517
 Db 2040 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 2099
 QY 1518 GTTTTATGATTCATCAACAGATCCATGAAAGAAAGAGCAAGCAAGGAGATTT 1577
 Db 2100 GAGCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 2159
 QY 1578 CGAGATTTGACAGCAGAGAACTGCAAGGTTGTTGACAGCAGCAGAGCAATTA 1637
 Db 2160 GGAATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGA 2219
 QY 1638 TCCCTTAGAATGACGATTCGCAAGAGAGCTGAGAAAGTCAAGCTTCATCGATT 1697
 Db 2220 GCAGCAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 2279
 QY 1698 TCAAGAAAGAGATGAGAGATTTTGAAGAGAGAGAGATGCTGTAAGATCAGA 1757
 Db 2280 GCAGCAGAGAGAGAGAGAGAGATTTAGAGAGCAGAGAGAGATTTAGAGATCAGA 2339
 QY 1758 GAAGAATGGAAGACATGAAGAGAGATCAAGAGAGATTTATGATCTGAGAAAGA 1817

Db 2340 GCAGAGATTAGAGAGCAGAGAGATTTAGAGAGCAGAGAGATTAGAGAGCA 2399
 QY 1818 ATTGATGAGGCTTTGGAAACG 1839
 Db 2400 GGAGCAGAGATTAGAGAGCAG 2421

RESULT 7

US-08-770-379-20/C
 ; Sequence 20, Application US/08770379
 ; Patent No. 5849564
 ; GENERAL INFORMATION:
 ; APPLICANT: Chang, Yuan
 ; APPLICANT: Bohenzky, Roy A.
 ; APPLICANT: Russo, James J.
 ; APPLICANT: Edelman, Isidore S.
 ; TITLE OF INVENTION: POLYPEPTIDES FROM KAPOSI'S SARCOMA-ASSOCIATED
 ; TITLE OF INVENTION: HERPESVIRUS, DNA ENCODING SAME AND USES THEREOF
 ; NUMBER OF SEQUENCES: 20
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSES: Cooper & Dunham LLP
 ; STREET: 1185 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: U.S.A.
 ; ZIP: 10036
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: PatentIn Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/08/770,379
 ; FILING DATE:
 ; CLASSIFICATION: 435
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: White, John P.
 ; REGISTRATION NUMBER: 28,678
 ; TELEPHONE/DOCKET NUMBER: 52342
 ; TELEPHONE: (212) 278-0400
 ; TELEFAX: (212) 391-0525
 ; INFORMATION FOR SEQ ID NO: 20:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 32207 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: DNA (genomic)
 US-08-770-379-20

Query Match 3.1%; Score 58.8; DB 2; Length 32207;

Best Local Similarity 43.4%; Pred. No. 2.5e-06; Mismatches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

QY 1218 GAAGCAAGATCTGACATATTCATCACTCTCAAGGCAAAACAGCTGAAATTCGA 1277
 Db 20197 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 20138
 QY 1278 GTTGAATCATACCAAGATGTTGTTAAAGAGCTGAGGAGATCTTGAGACAATCA 1337
 Db 20137 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 20078
 QY 1338 GCAGCTGAATCTTTAAGAACAAAGCTCTCAAAACAGAAACAGCCAGGCTTGA 1397
 Db 20077 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 20018
 QY 1398 GGAATCTCGAAATTTATGAGAGAGAGCTGCTGAATGACAGAGATTAATCGATCGT 1457
 Db 20017 GCAGCAGATGACAGCAGAGATGACAGCAGAGATGACAGCAGAGATGACGA 19958

QY 1458 GAGACAGAGAACTAAGATGACGATGAAACAGAAACAGGAAAGATGATGACACAGCAG 1517
 Db 19957 GCAGCAGATGACGACAGCAGATGATGACGACGATGACGACGATGACGACGATGACG 19898
 QY 1518 GTTTTCATGATTCATCAATCAATCCATGAAAGAGAGAGAGAGAGAGAGATTT 1577
 Db 19897 GGAGCAGCAGATGACGACAGCAGATGATGACGACGATGACGACGATGACGACG 19838
 QY 1578 CGAGATGTTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 1637
 Db 19837 GATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGATGA 19778
 QY 1638 TCCCTTACGATGACGATGACGACGACGACGACGACGACGACGACGACGACGAT 1697
 Db 19777 GCAGCAGCAGCAGATGACGACGACGACGACGACGACGACGACGACGACGACG 19718
 QY 1698 TCAGAGAGAGAGATGAGAGATGTTGTGAGAGAGAGAGATCTGATTAAGATCAAGA 1757
 Db 19717 GCAGCAGAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGATGAGAGATCA 19658
 QY 1758 GAAGAGATGAGAGAGATGAGAGAGAGATGAGAGAGATGAGAGATGAGAGAGAT 1817
 Db 19657 GCAGGAGTTAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGAGATGAGAGAG 19598
 QY 1818 ATTTGATGAGGCTTTGGAACAG 1839
 Db 19597 GGAGCAGAGATGAGAGAGAGAG 19576

RESULT 8
 US-08-757-669A-20/c
 ; Sequence 20, Application US/08757669A
 ; Patent No. 6183751

GENERAL INFORMATION:

APPLICANT: Chang, Yuan
 APPLICANT: Bohenzky, Roy A.
 APPLICANT: Russo, James J.
 APPLICANT: Edelman, Isidore S.
 APPLICANT: Moore, Patrick S.
 TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS
 NUMBER OF SEQUENCES: 20
 CORRESPONDENCE ADDRESSES:
 ADDRESSEE: Cooper & Dunham LLP
 STREET: 1185 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: U.S.A.
 ZIP: 10036
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patentin Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/757,669A
 FILING DATE:

CLASSIFICATION: 424

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.
 REGISTRATION NUMBER: 28,678
 REFERENCE/DOCKET NUMBER: 45185-F
 TELEPHONE: (212) 278-0400
 TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:
 LENGTH: 32207 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: DNA (genomic)
 US-08-757-669A-20

Query Match 3.1%; Score 58.8; DB 3; Length 32207;
 Best Local Similarity 43.4%; Pred. No. 2.5e-06;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

QY 1218 GAAGCAGATCTGACATATTCATCAATCACTCTCAAGCAAAACAGCTGAATTCGA 1277
 Db 20197 GCAGCAGATGACGACGACGACGACGATGACGACGATGACGACGATGACGACG 20138
 QY 1278 GTTGAATCATACCAAGATGTTGTAAAGAGCTGAAGCAGATCTGAGACATCA 1337
 Db 20137 GCAGCAGATGACGACGACGACGATGACGACGATGACGACGATGACGACGATG 20078
 QY 1338 GCAGTGAATCTCTTAAAGAACTCTCAAAACAGAACAGACGACGACGATGCTTGA 1397
 Db 20077 GCAGCAGATGACGACGACGACGATGACGACGATGACGACGATGACGACGATG 20018
 QY 1398 GGAATCTCTGAAATTTATGAGCGAGAGAGCTGCTGTAACCTGACAGATTAATCG 1457
 Db 20017 GCAGCAGATGACGACGACGACGATGACGACGATGACGACGATGACGACGATG 19958
 QY 1458 GAGACAGAGAACTAAGATGACGATGAAACAGAAACAGGAAAGATGACACAGCAG 1517
 Db 19957 GCAGCAGATGACGACGACGACGATGACGACGATGACGACGATGACGACGATG 19898
 QY 1518 GTTTTCATGATTCATCAATCAATCAATCCATGAAAGAGAGAGAGATTT 1577
 Db 19897 GCAGCAGAGATGACGACGACGACGACGATGACGACGATGACGACGATGACGAC 19838
 QY 1578 CGAGATGTTGACGACGACGACGACGACGACGACGACGACGACGACGACGACGAC 1637
 Db 19837 GATGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACGACG 19778
 QY 1638 TCCCTTACGATGACGATGACGACGACGACGACGACGACGACGACGACGACGAT 1697
 Db 19777 GCAGCAGCAGCAGATGACGACGACGACGACGACGACGACGACGACGACGACG 19718
 QY 1698 TCAGAGAGAGAGATGAGAGATGTTGTGAGAGAGAGAGATCTGATTAAGATCAAGA 1757
 Db 19717 GCAGCAGAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGATGAGAGATCAAGA 19658
 QY 1758 GAAGAGATGAGAGAGATGAGAGAGAGATGAGAGAGATGAGAGATGAGAGAGAT 1817
 Db 19657 GCAGGAGTTAGAGAGAGAGAGAGATGAGAGAGAGAGAGAGATGAGAGAGAG 19598
 QY 1818 ATTTGATGAGGCTTTGGAACAG 1839
 Db 19597 GGAGCAGAGATGAGAGAGAGAG 19576

RESULT 9

US-09-230-371A-20/c
 ; Sequence 20, Application US/09230371A
 ; Patent No. 6348586

GENERAL INFORMATION:

APPLICANT: Chang, Yuan
 APPLICANT: Bohenzky, Roy A.
 APPLICANT: Russo, James J.
 APPLICANT: Edelman, Isidore S.
 APPLICANT: Moore, Patrick S.
 TITLE OF INVENTION: UNIQUE ASSOCIATED KAPOSI'S SARCOMA VIRUS SEQUENCES AND
 FILE REFERENCE: 45185-G-PCT-US
 CURRENT APPLICATION NUMBER: US/09/230,371A
 CURRENT FILING DATE: 1999-11-17
 PRIOR APPLICATION NUMBER: PCT/US97/13346
 PRIOR FILING DATE: 1997-07-22
 NUMBER OF SEQ ID NOS: 30
 SOFTWARE: Patentin Ver. 2.0
 SEQ ID NO 20

LENGTH: 32207
 TYPE: DNA
 ORGANISM: Kaposi's sarcoma-associated herpesvirus

US-09-230-371A-20

Query Match 3.1%; Score 58.8; DB 4; Length 32207;
 Best Local Similarity 43.4%; Pred. No. 2.5e-06;
 Matches 270; Conservative 0; Mismatches 352; Indels 0; Gaps 0;

1218 GAAGCAAGATTCGACATATTCATCACTCTCAAGCAAAACAAAGCTGAATTGCA 1277
 20197 GCAGCAGGATGACAGCAGCAGAGATGAGCAGCAGAGTGGACGACGAGATGAGCA 20138
 1278 GTTGAATCATACCAAGAGTGGTGTAAAGAGCTGAGGCAATCTCTGAGCAATCA 1337
 20137 GCAGCAGGATGACAGCAGCAGATGAGCAGCAGAGTGGACGACGAGATGAGCA 20078
 1338 GCAGCTGAATCTTTAAAGAAAGCTCTCAAAACAAAGCAAGCCAGGCTTGA 1397
 20077 GCAGCAGGATGAGCAGCAGAGATGAGCAGCAGAGTGGACGACGAGATGAGCA 20018
 1398 GGAATCTCTGGAATTTTGAAGCGAAGCTGCTAGAACTGCAAGATTAATCGATCGT 1457
 20017 GCAGCAGGATGAGCAGCAGAGATGAGCAGCAGAGTGGACGACGAGATGAGCA 19958
 1458 GAGACAGAACTAAGATGAGCAGTGAACAGAAAGGAAAGATGATGACACAGCAG 1517
 19957 GCAGCAGGATGAGCAGCAGAGATGAGCAGCAGAGTGGACGACGAGATGAGCA 19898
 1518 GTTTCATGATTCATCAACAGATCCATGAAGAAAGAGCAAGAGAGGAAATT 1577
 19897 GCAGCAGGATGAGCAGCAGAGATGAGCAGCAGAGTGGACGACGAGATGAGCA 19838
 1578 CGAAGTGTGAGCAGCAGAACTGCCAAGTGTGTCGACAGCAGCAGCAATTA 1637
 19837 GATATGACAGCAGCAGAGTGGACGACAGCAGAGTGGACGACGAGATGAGCA 19778
 1638 TCCCTCTGCAATGACATTCGCCAAGAGCTGAGGAAGTCCAGCTTCACTGAGT 1697
 19777 GCAGCAGGATGAGCAGCAGAGATGAGCAGCAGAGTGGACGACGAGATGAGCA 19718
 1698 TCAAGAGAAAGATGAGAGATTTGTGAAAGAGAGAGATGCTGATTAAGATCAAG 1757
 19717 GCAGCAGGATGAGCAGCAGAGATGAGCAGCAGAGTGGACGACGAGATGAGCA 19658
 1758 GAAGAGATGAGAGATGAGAGAGCAGATCAAGAGATTTGATCTGAGAAAGA 1817
 19657 GCAGCAGGATGAGAGAGCAGAGAGATTTAGAGAGCAGAGAGATTTAGAGAGCA 19598
 1818 ATTGATGAGGCTTTGAAAG 1839
 19597 GCAGCAGGATTTAGAGAGCAG 19576

RESULT 10

US-08-781-891-208/c

Sequence 208, Application US/08781891

Patent No. 6090620

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-En

APPLICANT: Oshima, Junko

APPLICANT: Mulligan, John T.

APPLICANT: Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESS: SEED and BERRY LLP

STREET: 6300 Columbia Center, 701 Fifth Avenue

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/781,891
 FILING DATE: 27-DEC-1996
 CLASSIFICATION: B00
 ATTORNEY/AGENT INFORMATION:
 NAME: No. 6090620, endurg, Ph.D., Carol
 REGISTRATION NUMBER: 39,317
 REFERENCE/DOCKET NUMBER: 240052.419
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (206) 622-4900
 TELEFAX: (206) 682-6031
 INFORMATION FOR SEQ. ID NO. 208:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 16442 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: single
 TOPOLOGY: linear

US-08-781-891-208

Query Match 2.8%; Score 53.2; DB 3; Length 16442;
 Best Local Similarity 47.6%; Pred. No. 7.6e-05;
 Matches 157; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

1472 AGATGAGCATGAACAGAAACAGGAAAGATGATGATGATGATGATGATGATGATGAT 1531
 16437 AGAGAGAGAGCAGAGCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 16378
 1532 CAATCAACAGATTCATGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1591
 16377 AGAGAGAGAGCAG 16318
 1532 AGAGAGAGAGCAG 16258
 16317 AGAGAGAGAGCAG 16258
 1652 ACATTTGCGAAG 1711
 16257 AG 16198
 1712 TGAAGAGATTTGTGAAG 1771
 16197 AG 16138
 1772 ACATGAAG 1801
 16137 AG 16108

RESULT 11

US-09-618-166-208/c

Sequence 208, Application US/09618166

Patent No. 6583112

GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui

APPLICANT: Yu, Chang-En

APPLICANT: Oshima, Junko

APPLICANT: Mulligan, John T.

APPLICANT: Schellenberg, Gerald D.

TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESS: Seed Intellectual Property Law Group

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 208:
SEQUENCE CHARACTERISTICS:
LENGTH: 16442 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 208:
US-09-618-166-208

Query Match 2.8%; Score 53.2; DB 4; Length 16442;
Best Local Similarity 47.6%; Pred. No. 7.6e-05;
Matches 157; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

QY 1472 AGATGAGCATGAAACGAAGGAGAGATGATGACACGACGATTTTCATGAT 1531
DB 16437 AGAGGAGGAGCGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGG 16378
QY 1532 CAATCAAGATCCATGAAAGAGAGAGAGAGAGAGAGAGAGATTTTCAGAT 1591
DB 16377 AGC 16318
QY 1592 AGCAG 1651
DB 16317 AGC 16258
QY 1652 AGCATTCGCGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1711
DB 16257 AG 16198
QY 1712 TGGAG 1771
DB 16197 AG 16138
QY 1772 ACATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATAT 1801
DB 16137 AGAT 16108

RESULT 12
US-08-781-891-209/c
Sequence 209, Application US/08781891
Patent No. 6090620
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/781,891
FILING DATE: 27-DEC-1996
CLASSIFICATION: 800
ATTORNEY/AGENT INFORMATION:
NAME: No. 6090620cemburg Ph.D., Carol
REGISTRATION NUMBER: 39,317
REFERENCE/DOCKET NUMBER: 240052.419
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
US-08-781-891-209

Query Match 2.7%; Score 49.8; DB 3; Length 51259;
Best Local Similarity 50.5%; Pred. No. 0.0015;
Matches 146; Conservative 0; Mismatches 142; Indels 1; Gaps 1;

QY 1525 ATGATTCATCAATCAACATCCATGAAAGAGA-GAGCGAAGAGAGAGATTTTCAGAT 1583
DB 290 ATGCTTAAACACACAAATTAATCATGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 231
QY 1584 GTTGCGAGCAG 1643
DB 230 GAG 171
QY 1644 TAGCAATGACATTTGCCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1703
DB 170 GCAG 111
QY 1704 GAAG 1763
DB 110 GAG 51
QY 1764 GATGGAAGCATGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGATTTGATCTGAG 1812
DB 50 GAG 2

RESULT 13
US-09-618-166-209/c
Sequence 209, Application US/09618166
Patent No. 6583112
GENERAL INFORMATION:
APPLICANT: Fu, Ying-Hui
APPLICANT: Yu, Chang-En
APPLICANT: Oshima, Junko
APPLICANT: Mulligan, John T.
APPLICANT: Schellenberg, Gerald D.
TITLE OF INVENTION: GENE AND GENE PRODUCTS RELATED TO
TITLE OF INVENTION: WERNER'S SYNDROME
NUMBER OF SEQUENCES: 209
CORRESPONDENCE ADDRESS:
ADDRESSER: Seed Intellectual Property Law Group
STREET: 701 Fifth Avenue, Suite 6300
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104-7092
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30

```

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/618,166
FILING DATE: 17-Jul-2000
CLASSIFICATION: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: McMaister, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 240052.419C1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 209:
SEQUENCE CHARACTERISTICS:
LENGTH: 51259 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 209:
US-09-618-166-209

Query Match
Best Local Similarity 2.7%; Score 49.8; DB 4; Length 51259;
Best Local Similarity 50.5%; Pred. No. 0.0015; Indels 1; Gaps 1;
Matches 146; Conservative 0; Mismatches 142;

QY 1525 ATGATTCATCAACAGATCATGAAAGAGA-GAGCAAAGAGAGAGATTTCAGAT 1583
DB 290 ATGCTTAAACACAAATCATGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 231
QY 1584 GTTCAGACAGAGAAAGCTGCCAAGTTTGTGGCCACAGACAGAGAACTTATCCCTC 1643
DB 230 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 171
QY 1644 TAGCAATACATTCGCCAAGAGAGCTGAGAGAGTCTCAAGCTTCAGATTCAAG 1703
DB 170 GCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 111
QY 1704 GAAAGAGATGAGAGAGTTGTGAAAGAGAGAGAGATGCTGATAAAGATCAAGAGAG 1763
DB 110 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 51
QY 1764 GATGAGAGCATGAAAGAGAGCATCAGAGAGATATTGATCTGAG 1812
DB 50 GGAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 2

RESULT 14
US-09-177-325-2
Sequence 2, Application US/09177325B
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins,
FILE REFERENCE: LUD 5525
CURRENT APPLICATION NUMBER: US/09/177,325B
CURRENT FILING DATE: 1998-10-22
NUMBER OF SEQ ID NOS: 15
SEQ ID NO 2
LENGTH: 1276
TYPE: DNA
ORGANISM: Homo sapiens
FEATURES:
Query Match
Best Local Similarity 2.5%; Score 47.4; DB 3; Length 1276;
Best Local Similarity 47.2%; Pred. No. 0.00088; Indels 0; Gaps 0;
Matches 144; Conservative 0; Mismatches 161;

QY 1516 AGTTTTCATGATTCATCAATCAACAGATCCATGAAAGAGAGCGCAAGAGAGCAAT 1575
DB 607 ATGTCTTAAAGAAATCATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 1576 TTGAGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1635
DB 667 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726
QY 1636 AATCCCTTAGCAATGAGATTTGCCAAGAGAGAGAGAGAGAGAGAGAGAG 1695
DB 727 TAATAGTAGTAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
QY 1696 TTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755
DB 787 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
QY 1756 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
DB 847 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906
QY 1816 GAATT 1820
DB 907 AATT 911
```

```

RESULT 15
US-09-411-812A-2
Sequence 2, Application US/09411812A
GENERAL INFORMATION:
APPLICANT: Tureci, Ozlem
APPLICANT: Sahin, Ugur
APPLICANT: Pfreundschuh, Michael
TITLE OF INVENTION: Isolated Nucleic Acid Molecules Which Encode SCP Proteins, And
FILE REFERENCE: LUD 5525.1 CIP
CURRENT APPLICATION NUMBER: US/09/411,812A
CURRENT FILING DATE: 1999-10-01
PRIOR APPLICATION NUMBER: US 09/177,325
NUMBER OF SEQ ID NOS: 14
SEQ ID NO 2
LENGTH: 1276
TYPE: DNA
ORGANISM: Homo sapiens
US-09-411-812A-2

Query Match
Best Local Similarity 2.5%; Score 47.4; DB 3; Length 1276;
Best Local Similarity 47.2%; Pred. No. 0.00088; Indels 0; Gaps 0;
Matches 144; Conservative 0; Mismatches 161;

QY 1516 AGTTTTCATGATTCATCAATCAACAGATCCATGAAAGAGAGCGCAAGAGAGCAAT 1575
DB 607 ATGTCTTAAAGAAATCATCGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 666
QY 1576 TTGAGATGTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1635
DB 667 CCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 726
QY 1636 AATCCCTTAGCAATGAGATTTGCCAAGAGAGAGAGAGAGAGAGAGAGAG 1695
DB 727 TAATAGTAGTAGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 786
QY 1696 TTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1755
DB 787 AGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 846
QY 1756 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1815
DB 847 GAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 906
QY 1816 GAATT 1820
DB 907 AATT 911
```

Fri Apr 9 15:36:15 2004

Search completed: April 8, 2004, 16:29:41
job time : 96.7349 secs

us-10-030-829-2.rml

Page 9

This Page Blank (uspto)


```

Db      572 TTTTGAGACTCGAAAAACGAACAGTGTCTCAAGAGTTCTTTGAGATATGAAATCATTT 631
Qy      642 GTGCAATCGACATATTAATTAACACAGAGGCACTGTGCTATGTCAGCTTGTCAAGACG 701
Db      632 GAGTTTGGAACAAATTAATTAACAACTCGGCAATGTGCTATGTCGCGCATGCAAAAATG 691
Qy      702 ACCGTGTCATGATTTGTATTA---CCTGACCCCTCTACTAGTCTATGCGAGACAA 758
Db      692 ACCGTGTCATGATTTGTATTA---CCTGACCCCTCTACTAGTCTATGCGAGACAA 751
Qy      759 AGGAGCTAGGCGAGTTAAGCTCCATAGAGAAATTTGCTGAATTTTGAAGAGATCTACA 818
Db      752 GGGTTCTACAAAGGTTAACTTACAGAGAAATTTGCTGATCTCTGAGAGAGAGCTTTC 811
Qy      819 GATGAAAGGCGCATCTGTCTATCTCTGTGTGAGATTTATGCGCAGTGAAGGTTTGG 878
Db      812 TCGCAGGGAACCTTCAGTGTACACGCGGTGAACAAATTTGGAAATGGAAGAGCTG-- 869
Qy      879 TGAGATGAAGAAAGATTAATTAATTTGCTGCTCCATGTCATCATCATCAATCAATAG 938
Db      870 -CAAGAAAGACATGATCGTGAAGATAGTGTGCGCACCAATGTATTTATGATACCTT 928
Qy      939 ACTGATTAAGAGACATTAACGATTAAGTGTGCTGCGCATGCGCAACCAAGCTGCTGAATA 998
Db      929 CTGGAAGAAAGATGAAGATATTAAGTGAAGGCGCATGCGAAACCAAGAGCTCTTGATTA 988
Qy      999 CTTCGACAAATAGAGCTCTTAGAGCAGCGCCATTTCTATGTGTCACAGAGGCTATGTG 1058
Db      989 TTTTGAGAAATTAAGAACCAATTAACGATTAAGTGTGCTGCTGCGGACCGTGTG 1048
Qy      1059 GATGAGTGTCTGATGTTTGAAGAGAGTGCCTGCTGCTATTTTGGAGCGCAACGCTTCA 1118
Db      1049 TATGACGCTGTTAATTTGAAGAGCTCAGCTGTGCGCTATCATGAGAGCTGAAGCTTGA 1108
Qy      1119 CCGGAGTTAGCTGAGATGGGTTAAGATTAAGTGTGCTGCGGCTCA--GAAGCGAGTAT 1175
Db      1109 TAAACATTTGTTAATCAAGGTACAGAGATTAATCATGCGACCTACGCAAGGTTGAT 1168
Qy      1176 GTTTTGTGAGAGTGTGCGCAACTGTATGTGCTTCTTGAAGACAGACAGATTTGACAT 1235
Db      1169 TGTGCTGTGGGAAAGGCACTAATATATGTTCTTGAACAAAGAGAGATTAAGAGGC 1228
Qy      1236 ATTCATTAACAACCTCAAGGCAAAACAAGGCTGAATTTGAGTGAATCATACCAAGA 1295
Db      1229 ATTTAACAAGATTTGCAATGTGTAAGGCGCTGTAAGTATGAGATGAGGCTTACATGA 1288
Qy      1296 GATGTTTGAAGAGAGCTGAGAGATCTTCTGAAGCAATCAGCAGCTGAATCTTTAA 1355
Db      1289 GATGTTTGAAGAGAGCTGAGAGATCTTCTGAAGCAATCAGCAGCTGAATCTTTAA 1348
Qy      1356 GAAACAGCTCTCAAAACAGAACAGACCGCAAGTGTGTTGAGAGATCTCTGAAATTTAT 1415
Db      1349 GAAACAGATGTGTTAACAAGAGCAGATCTTAAAGAGTAAAGAAATCAATGGGTGTGT 1408
Qy      1416 GAGCGAAGAGCTGCTGAAGATCTGAGAGATTAATCGATCTGTGAGACAGAGAACTAAGAT 1475
Db      1409 TATCCGAAGAGCTACGTGAAGCTATTAAGAAATTAATTTGTTAGAGATTAAGCAAAA 1468
Qy      1476 GCAGCATGAACAGACAGAGAGATGATGACACAGCAGGTTTTTCATGTGATTTCAAT 1535
Db      1469 GAGGACATGAGATGAGAGAGATGAAGTCCAGAGAGAAATTTTTCATGTCTAAT 1528
Qy      1536 CAAACAGATCTCAATGAAGAAAGACCAAGAGAGAGAAATTTGCAATGTGTGAGAGCA 1595
Db      1529 TGAAGATTAATTCAGAGCCACAGAGATTAAGAAAGAGTTTGAAGATTTGTGAGGA 1588
Qy      1596 GGAACGTGCAAGAGTTGTGCGCAGAGCAGAGACATTAATCCTCTAGCAATGACGA 1655
Db      1589 AAGAGCCTCAAGAGCCAGAGC-----TTTGAATGTGATTTCTGGAACATTAAGAA 1639
Qy      1656 TTGCCGAAAGAGAGCTGAGAGATGTCAAGCTTCATGAGTTTCAAGAGAAAGAGATGA 1715

```

```

Db      1640 TCGCAGCTAAGAAAGAAATATGTGCAGAAAGTTTCATGTATGTCACAGTCAAGATGTGCG 1699
Qy      1716 GGAAGTTTGTGAAGAGAGAGAAATCTGATTAAGATCAAGAAAGAAATGGAAGACT 1775
Db      1700 AAGATTTGAAGTTGAGCGAGAGAGAGCTGATCAAAATCCACGAGGACAAAGAGTTGAAGCT 1759
Qy      1776 GAAGAAAGAGCATCAAGAGAGATATTTGATCTGGAAGAAAGAAATTTGAGAGCTTTTGA 1835
Db      1760 CAAGAAAGAAATCAATGATCTGTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1819
Qy      1836 ACAGCTCATGTACAAAGCAT 1854
Db      1820 GGGCTGATGAGAAAGCAT 1838

RESULT 2
US-10-424-599-90522/c
; Sequence 90522, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 90522
; LENGTH: 2218
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(2218)
; OTHER INFORMATION: unsure at all n locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_52754C.1
US-10-424-599-90522

Query Match      21.3%; Score 400.4; DB 12; Length 2218;
Best Local Similarity 67.1%; Pred. No. 1.3e-105;
Matches 615; Conservative 0; Mismatches 291; Indels 10; Gaps 3;

Qy      455 AGGAGTTTCTGACGTGAGAGATGATGTGATATGCTTCTGAGAGAGAAATGATTCGG 514
Db      2153 AAGAGTCCCTTATGAAGATGATGATGATGATGATGAAGAGGAGTGAAGAGAGGTTTG 2094
Qy      515 ATGCTTTGGATGAT-TCTGATGAGACGCTTGAAGTGAATGATGATCGATCGATGATGAT 573
Db      2093 ATGCTATGAGAGATACCGAGATGATGATCTAATGATGATGATGATGATGATGATGATGAT 2034
Qy      574 CAAAGAGCAGATGATCAAGAAAGAGATTAAGTGTTCAAAGATCTTTGGCAGCTTG 633
Db      2033 CAAAGAGACAGAGACTGTGAAGAGAGCAAAATGTTTGAAGATTTCTTTGAGAGACTTG 1974
Qy      634 GATAGCTGTGATGATGAGAGATTAATTAACACAGAGCAGTGGCATTTGTCCAGCTTGT 693
Db      1973 GATGCTTATCTATGAGAAAGATCAATGAACCGAAAGGAGGAGCTCTGATTAATCTATGTC 1914
Qy      694 CAGAAAGGACCTGTGCGCATGATTTGTGATA--ACCTGCAACCTCTACTAGCTGATGCG 750
Db      1913 CGAGGTGTGCTCCGGGTGATTTGATTTGTGACAGAGAGCTGAGCTCTGATTAATCTATGTC 1854
Qy      751 AGGACAAAGAGAGCTAAGGCGAGTTAAGCTCCATAGAGAAATTTGCTGAAGTTTGAAGAA 810
Db      1853 AAAACAAAGGAGTCAAAAGAGGATGAAGATCCATAGAGAGTTTGTGAGAGAG 1794
Qy      811 GATCTACAGATGAGAGGCGCATCTGTCAATCTTGTGTGATGATTTATGCGAGTGAAG 870
Db      1793 GAACTGCCAGAGAGGCGCTGTGCAATATTCACCTGGGGAAGATTTTGTGAAGTGAAG 1734

```


QY	871	GGTTTGGGTGAGGATGAAAGAGATTAAGAAATTGCTGGCCCTCCAAATGGTCATCATG	930
Db	1733	GGTTTAAAGATGAGAGAAAGATCATGAAATAGTTGGCCCTCCAGATGGTGTCAATCAG	1674
QY	931	AATACTAGACATCGATTAAGAGACATTAACGATTAAGTGGCTCGGCATAGGGCAACCAAGACTG	990
Db	1673	AATACAAAGCTTGGAAACGAGATGAAATATGATAGTGGCTAGGTATAGGTAAACAGAGACTT	1614
QY	991	CTGGAATACCTTCGACAAAGTATAGAGCTCTTAAGACCGCCATTCCTATGTTCCAAGGGC	1056
Db	1613	CTTAATATATTTTAAAGCATATGATGCTGTGAAGAGCTGCACACGGCTTAATGGCCCCCAGGAT	1554
QY	1051	CATGCTGGAGATGATGTTCTGATGTTTAAAGACAGTGCACATGGCTATTTTGAAGGCCAA	1110
Db	1553	CATGGGGGTTGAACGTTTATATTGTAATTTGAACATCAGTATAGTTATCTTGAAGGCTGAG	1494
QY	1111	CGCCCTCAACGGGAGTTAGCTGAGATGGGGTTAGTAAATATGCTCGGGGTCAAGAGGC	1170
Db	1493	CGTCTAACCAAGCACTTTGCAGAACAAAGAACTGACGAGATGCTTGGTTTATGTCACAC	1434
QY	1171	AGTATGTTTT-----CTGAGAGTGTTCGCCAACTGTATGGCTTCTTGCACAGAAAGCA	1224
Db	1433	CGTAGATTATTTCTCCCTGCTGGCCGGGACAGACTCTATGATATCATGCAATTAABAA	1374
QY	1225	GATCTGACATATTTCAATCAACACTCTCAAGGCAAAACAAAGGCTGAAATTCGATGTAA	1284
Db	1373	GACCTGGACTTTTTCACACAGACATTTGCCAAAGTAAATCTAGACTCAATATATGACATGAGA	1314
QY	1285	TCATTAACCAAGATGCTGTAAAGAGCTGAGGACAGATCTCTGAGAGCAATTCAGACCTG	1344
Db	1313	TCATTAACCAAGATGCTGTAAACCACTTGGCAAAAGAAATGAGGCCAACCAAGCTGCT	1254
QY	1345	AACACTTAAAGACA	1360
Db	1253	AATATATTTGAAGAAA	1238

```

RESULT 3
US-10-424-599-91177
Sequence 91177, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 91177
LENGTH: 1377
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_53343C.1
US-10-424-599-91177

```

	Query Match	13.6%	Score 254.6	DB 12	Length 1377
	Best Local Similarity	66.3%	Pred. No. 3.5e-63		
	Matches 417	Conservative 0	Mismatches 199	Indels 13	Gaps 3
Qy	352	TATGATTAACATTGTTGCGACACCCACCTGATCTGCCCTCTTGGAGAGAGATGG	411		
Db	29	TATGAAACCACTATGTACTTCCAAATCCTGTAAATGTGGCCTCCACTGAGCATGGCTGG	88		
Qy	412	AATGGCGAGGAGAGAGAGATTTCTGCTCAGACACAGCTGTGCGAGAGATTTCTTGACGTG	471		
Db	89	AACCTGGCATCCAGACTGTGTGCAATTATGCTCCATGTGAAGGATGAATAATTCACCAAG	148		

Oy	472	G-----	AGATGATGATGATGATTAATGCTTCTGAGGAAGAAATATTCCTGATGCTTTG	522
Db	149	GACCTTCAAAAAGAAATTATG	TGTGTGATGATGAATGAGGAGGAGGAATTCATGATTTTG	208
Oy	523	GATGATTTCTGATGACGACCTT	GCAAGTATGATTTATGACTCGATGTGAGTCAAAAAGAC	582
Db	209	GAAATATCTATGATGATGAT	CTGATGATGATGACGATGATTTCTGATCTGATGATCAAAAAGAT	268
Oy	583	CATGATTCACGAAAGCAGAA	TAAAGTGTTCAAAAAGTTCTTTGGACAGCTTGATACCTTG	642
Db	269	CATGACACAGTAAAGAAAAG	CAATATGTAAGAGTTCTTTGATGATTTTGGATGGCTTG	328
Oy	643	TCGATCGACGATTAATTA	TGAACCAAGGACAGTGGATTTGTCCAGTTTCTCAGAACGA	702
Db	329	ACTGTTTAAACGATTAAC	GAAACGAAAGCAAGTGGATTTGTCCAGTTTCCCAAAGTGCT	388
Oy	703	CCTGATGACCATCGATG	TGATTAAC---CTGCACCTCTACTAGCTCACTGGAGGACAAAA	759
Db	389	CTTGAGACTATTGATTTGT	ATCAGAGCATCCAGCCCTTGATGACACATGCCAAAAACAAAG	448
Oy	760	GGAGCTAGGCGAGTTAA	AGCTTCATAGAGAAATGGCTGAAATTTTGAAGAAAGATCTACAG	819
Db	449	GGATCAAAAAGGGGTGA	AGATTCCACAGGAGACTTGCTATATCTTTTGGATGAGGAATTAACG	508
Oy	820	ATGAGAGCGCATCTG	ATCTCTTGTTGATGATTTATGAGGACAGTGAAGGGTTGGGT	879
Db	509	AAAAGGGGTACTTCAG	TAAATTTCCAGCTGCGGGAAGTATTTGAAAAGTGGAAAAGGTTTGAAA	568
Oy	880	GAGATGAAAAAGATTT	TGAAATTTGTCTGGCCCTCCAAATGTCATCATCAGTAATCTAGA	939
Db	569	GAAAGAG-AAAAGG	CCATGAGATAGTTTGGCTCTCCAAATGATGATTCAGAAATCAAAAG	627
Oy	940	CTGGATTAAGGACGAT	AACGATTAAGTGCT	968
Db	628	CTTGAAACGATGAAAA	TGATTAAGCGAT	656

```

, RESULT 4
, US-10-424-599-91176
, Sequence 91176, Application US/10424599
, Publication No. US20040031072A1
, GENERAL INFORMATION:
, APPLICANT: La Rosa Thomas J
, APPLICANT: Kovalick David K
, APPLICANT: Zhou Yihua
, APPLICANT: Cao Yongwei
, TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
, TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
, FILE REFERENCE: 38-21(53223)B
, CURRENT APPLICATION NUMBER: US/10/424,599
, CURRENT FILING DATE: 2003-04-28
, NUMBER OF SEQ ID NOS: 285684
, SEQ ID NO 91176
, LENGTH: 1589
, TYPE: DNA
, ORGANISM: Glycine max
, FEATURE:
, OTHER INFORMATION: Clone ID: PAT_MRT3847_53342C.1
, US-10-424-599-91176

```

Query Match	Best Local Similarity	11.5%;	Score 215.6;	DB 12;	Length 1589;
Matches	358;	Conservative	0;	Mismatches 104;	Indels 12; Gaps 2;
Qy	352	TATGATPACAACTTGTGTGCACACCCCACTATATCTGGCCCTCTTTGAAGAGAGATGG	411		
Db	29	TATGAAAGCACTATGTGATCTCCAAATCTGTATTTGGGCTTCCATCTGAGCATGGCTGG	88		
Qy	412	AATTGGCAGGCAAGAGAGGTTCTGCTCAGACACACAGCTGTGACAGAGTTTCCTCGACGTG	471		
Db	89	AACCTGGCATCCAAACCTGGTGCAATTCAGTCCAAATGTAAAGGATGAAATTTACACAGAG	148		
Qy	472	G-----AGATGTATGTGGATTAATGCTTTCTGAGGAGAGAAATGATTCGATGCTTTGG	522		

```

Db      149 GACCTTCAAAAAGATTATGTTGATGATGAGGAGGAGAGGAACTTGATGATTTG 208
Qy      523 GATGATTCGATGACGACCTTGCAAGTATGATTAATGCTGGATGATGATCAAAAAGAC 582
Db      209 GAAATACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 268
Qy      583 CATGATCAAGAAAGCAATTAAGTGTTCAAAAGTTCTTTGGAGCTGGATGATGATG 642
Db      269 CATGACACATGTAAGAAAGCAATGATTAAGTGTTCAAAAGTTCTTTGGAGCTGG 328
Qy      643 TCATGACACATGTAAGTAAATGAAACAGAGGAGGATGATGATGATGATGATGATG 702
Db      329 ACTGTTGAACATGTAAGTAAATGAAACAGAGGAGGATGATGATGATGATGATG 388
Qy      703 CCGTGCCATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 759
Db      389 CTGGAGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 448
Qy      760 GAGACTAGGCGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 819
Db      449 GATCAAAAAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 508
Qy      820 ATGAGAGGCGCATGATGATGATGATGATGATGATGATGATGATGATGATG 879
Db      509 AGAAGGGGCTCTGCAATGATGATGATGATGATGATGATGATGATGATGATG 568
Qy      880 GAGATGAAAAGCA 893
Db      569 GTGACGACGAAAAGA 582

```

RESULT 5
US-10-424-599-73594/C

```

; Sequence 73594, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 73594
; LENGTH: 2890
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37470C.1
US-10-424-599-73594

```

Query Match 10.5%; Score 197.8; DB 12; Length 2890;
Best Local Similarity 69.4%; Pred. No. 26-46;
Matches 284; Conservative 0; Mismatches 122; Indels 3; Gaps 1;

```

Qy      548 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 607
Db      2890 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2831
Qy      608 GGTTCAAAAGTTCTTTGGAGCTGGATGATGATGATGATGATGATGATGATGATG 667
Db      2830 GGTTCAAAAGTTCTTTGGAGCTGGATGATGATGATGATGATGATGATGATGATG 2771
Qy      668 AGAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 724
Db      2770 AAAGGAGGAGGATGATGATGATGATGATGATGATGATGATGATGATGATG 2711
Qy      725 ACCTGACCTCTTACTAGCTGATGATGATGATGATGATGATGATGATGATGATG 784

```

```

Db      2710 GACTGACCTCTGATTAATGATGATGATGATGATGATGATGATGATGATG 2651
Qy      785 GAGATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 844
Db      2650 GGAATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2591
Qy      845 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 904
Db      2590 GTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2531
Qy      905 TCTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 953
Db      2530 TTTGGCTTCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 2482

```

RESULT 6
US-10-424-599-36215/C

```

; Sequence 36215, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 36215
; LENGTH: 1103
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(1103)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_132704C.1
US-10-424-599-36215

```

Query Match 9.3%; Score 175.4; DB 12; Length 1103;
Best Local Similarity 65.7%; Pred. No. 3-9e-40;
Matches 285; Conservative 0; Mismatches 147; Indels 2; Gaps 2;

```

Qy      1172 GATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1231
Db      434 GTGGCTCTCTGAGGAGGATGATGATGATGATGATGATGATGATGATGATG 375
Qy      1232 ACATATTCATCAACACTCTGCAAGCAAAACAAAGCTGAAATTCGATGATGATG 1291
Db      374 ACTTTTCAACAGCACTCTGCAAGCAAAACAAAGCTGAAATTCGATGATGATG 315
Qy      1292 AAGAG-ATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1350
Db      314 AGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 255
Qy      1351 TTTAAGAACAGCTTCAAAACAGAACAGCAAGCAAGCTTGAAGATTCCTGANA 1410
Db      254 TTTAAGAACAGCTTCAAAACAGAACAGCAAGCAAGCTTGAAGATTCCTGANA 195
Qy      1411 ATTATGAGGAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 1469
Db      194 ATTATGAGGAGAGCTGCTGATGATGATGATGATGATGATGATGATGATGATG 135
Qy      1470 TAAAGTCAAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 1529
Db      134 TAAATCAAGTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 75
Qy      1530 TTTCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1589
Db      74 CCAATCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 15

```

QY 1590 GCAGCAGAACCTG 1603
 |||||
 Db 14 ACAGAGGAACGTG 1

RESULT 7

US-10-424-599-73591/C
 ; Sequence 73591, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; FILE REFERENCE: 38-21(53223) B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 73591
 ; LENGTH: 1375
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(1375)
 ; OTHER INFORMATION: unsure at all n locations
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_37468C.1
 ; US-10-424-599-73591

Query Match 9.0%; Score 169.8; DB 12; Length 1375;
 Best Local Similarity 69.0%; Pred. No. 1.9e-38;
 Matches 247; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

QY 548 GTGATATTATGCTCGATGTGATCAAAAGAGCCATGATACGAAACAGAAATTAAGT 607
 |||||
 Db 1375 GTATATATTATGATTTCTGATATCTGATCAAAAGAGCCATGATCGTAAGAAACAAAT 1316
 |||||
 QY 608 GGTTCAAAAAGTTCTTTGGCAGCTTGATAGCTTGTGATGACGAGTAATTAAGAAC 667
 |||||
 Db 1315 GGTTCAGAGATTTCTTTGAGAACTTGATGAGCTTATCTATCGAAAGATCAATGAACCGG 1256
 |||||
 QY 668 AGAGGAGTGGCATTTGTCCAGCTTGTGACAAAGAGAGCTGTGTCATGATTTGTTA--- 724
 |||||
 Db 1255 AAGGAGTGGCAGCCCTCAGCTTGTGAGGTGTCCTGCTATTTGATTTGATGACAGAG 1196
 |||||
 QY 725 ACCTGACCTCTACTAGCTCATGCGAGCAAAAGAGCTAGCGAGTTAAGTCCATA 784
 |||||
 Db 1195 GACTGACCTCTGATTTACTCATGCAAAAGAGGCTCAAAAGGCTGAAGATTCATA 1136
 |||||
 QY 785 GAGAAATTTGCTGAGTTTGAAGAAAGATCTACAGATGAGAGCGCATCTGTCATTCCTT 844
 |||||
 Db 1135 GGGAGTTTGTGAGCTTTTGAAGAGAGATGCGCAAAAGGGCTCTGCGATTAATTCAC 1076
 |||||
 QY 845 GTGGTAGATTTATGGCGAGTGAAGGCTTTGGCTGAGATGAAGAAAGATTAAGAAAT 902
 |||||
 Db 1075 CTGNGAAGATTTATTTGTAAGTGAAGGTTTAAAGATGAGGGGAAAGATCATGGAAT 1018
 |||||

RESULT 8

US-10-424-599-99770/C
 ; Sequence 99770, Application US/10424599
 ; Publication No. US20040031072A1
 ; GENERAL INFORMATION:
 ; APPLICANT: La Rosa Thomas J
 ; APPLICANT: Kovalic David K
 ; APPLICANT: Zhou Yihua
 ; APPLICANT: Cao Yongwei
 ; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
 ; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

FILE REFERENCE: 38-21(53223) B
 ; CURRENT APPLICATION NUMBER: US/10/424,599
 ; CURRENT FILING DATE: 2003-04-28
 ; NUMBER OF SEQ ID NOS: 285684
 ; SEQ ID NO 99770
 ; LENGTH: 951
 ; TYPE: DNA
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; OTHER INFORMATION: Clone ID: PAT_MRT3847_61108C.1
 ; US-10-424-599-99770

Query Match 5.8%; Score 109.6; DB 12; Length 951;
 Best Local Similarity 54.5%; Pred. No. 5.7e-21;
 Matches 313; Conservative 0; Mismatches 219; Indels 42; Gaps 3;

QY 1280 TGAATCATACCAAGATGTTGTGAAGAGCTGAGCAGATCTTGAGACATTCAGC 1339
 |||||
 Db 949 TGAGATCATATCAAGATGTTGTGAAGAGATTAACATTTATGATGATGATCGGA 890
 |||||
 QY 1340 AGCTGAATCTTTAGAACAGCTCTCAAAACAGAACAGACCGCAGAGTCTTGAG 1399
 |||||
 Db 889 AGCTGACTATTTATAGAGATGTTGACCAAGAACAAATCAATACAAATGCTGAG 830
 |||||
 QY 1400 AATCTTGAAATTTATGACGAGAGAGCTGTAAGATCTGAGAGATTAATCGATCGTA 1459
 |||||
 Db 829 ACTCTTATGTAATCAATGAGAAATTAAGATGACATGAACTGAAAAAACCGTGTGAC 770
 |||||
 QY 1460 GACAGAACTAAGATGACGATGAACAGAACAGGAGAGATGATGACACAGCAGT 1519
 |||||
 Db 769 AACAGACGAC-----AGAGATGATGACCTAAGAAAAA 737
 |||||
 QY 1520 TTTTCATGATTCATCAACAGATCCATGAAAGAGAGACGCAAGAGAGAAATTTG 1579
 |||||
 Db 736 ATTTCCAGAGCCAAATCCAGACATTCACAGACCTAGCTGCAAGAGAGATTAAGTTG 677
 |||||
 QY 1580 AGATGTTGACAGACAGAGAACTGCCAAGTTGTTGCCAGACAGCAGAACTTAATC 1639
 |||||
 Db 676 TGAATTTACAGCTGCAAG-----GCAATGCAAGAAAGATGAAGAGTCTCGTGAAGAAAT 623
 |||||
 QY 1640 CCTTCACATGACGATTCGCCAAGAGAGCTGAGAGAGTTCATGATGAGTTTC 1699
 |||||
 Db 622 CTTCGAGAAAGAGGAC-----AAGTGAAGAAATTCCTAGTTTCCTGAACCTTC 572
 |||||
 QY 1700 AAGAGAAAGATGAGAGAGTTTGTGAGAGAGGAGATGCTGAATAAAGATCAAGAGA 1759
 |||||
 Db 571 AAGACAGAGAGATGAAGCAATTCAGGACAGAGAGAGAAATTAATAAATTCATGAAG 512
 |||||
 QY 1760 AGAAGATGAGAGATGAAGAGAGCATCAGAGAGATATTGATCTGAGAAAGAAAT 1819
 |||||
 Db 511 AGAAAAAGTTGGCTCTGAGAAAGAGAGAGATGCGCAAGAGAGGTTGAATTTGAGAGAGAGT 452
 |||||
 QY 1820 TTGATGAGGCTTTGGAACAGCTCATGTACAGCA 1853
 |||||
 Db 451 TAGAGATGAATGACGACGATCATGACAGTA 418
 |||||

RESULT 9

US-09-294-093B-1296
 ; Sequence 1296, Application US/09294093B
 ; Patent No. US20010051335A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lalgudi, Raghunath, V.
 ; APPLICANT: Ito, Laura, Y.
 ; APPLICANT: Sherman, Bradley, K.
 ; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES DERIVED FROM CORN TASSEL
 ; FILE REFERENCE: PL-0009 US
 ; CURRENT APPLICATION NUMBER: US/09/294,093B
 ; CURRENT FILING DATE: 1999-04-16
 ; PRIOR APPLICATION NUMBER: 60/082,567
 ; PRIOR FILING DATE: April 21, 1998
 ; NUMBER OF SEQ ID NOS: 6207
 ; SOFTWARE: PERL Program

Sequence 208, Application US/10374077
Publication No. US20040006779A1
GENERAL INFORMATION:

APPLICANT: Fu, Ying-Hui
Oshima, Junko

Mulligan, John T.
Schellenberg, Gerald D.

TITLE OF INVENTION: ANTIBODIES AGAINST GENE PRODUCTS RELATED TO
WERNER'S SYNDROME

NUMBER OF SEQUENCES: 209

CORRESPONDENCE ADDRESS:

ADDRESSEE: Seed Intellectual Property Law Group

STREET: 701 Fifth Avenue, Suite 6300

CITY: Seattle

STATE: Washington

COUNTRY: USA

ZIP: 98104-7092

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentlin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/374,077

FILING DATE: 25-Feb-2003

CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Roseman, Stephen

REGISTRATION NUMBER: 43,058

REFERENCE/DOCKET NUMBER: 100107.401D1

TELECOMMUNICATION INFORMATION:

TELEPHONE: (206) 622-4900

TELEFAX: (206) 682-6031

INFORMATION FOR SEQ ID NO: 208:

SEQUENCE CHARACTERISTICS:

LENGTH: 16442 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

SEQUENCE DESCRIPTION: SEQ ID NO: 208:

US-10-374-077-208

Query Match

Best Local Similarity 47.6%; Score 53.2; DB 15; Length 16442;

Matches 157; Conservative 0; Mismatches 173; Indels 0; Gaps 0;

1472 AGATGACGATGAAACGACGAGGAGATGATGACACGACGATTTTTCATGATT 1531

16437 AGGAGGAGGAGCGAGCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16378

1532 CAATCAACGATTCATGAAAGAGAGAGCAAGAGAGGAGGAGGAGGAGGAGG 1591

16377 AGGAGGAGGAGCGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16318

1592 AGCAGAACTGCGCAAGTTGTTGGCCACGACGAGGAGGAGGAGGAGGAGG 1651

16317 AGGAGGAGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16258

1652 ACCATTCCGAAAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1711

16257 AGGAGGAGGAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16198

1712 TGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1771

16197 AGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 16138

1772 ACATGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1801

16137 AGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 16108

RESULT 15

US-10-108-260A-1966
Sequence 1966, Application US/10108260A
Publication No. US20040005560A1
GENERAL INFORMATION:

APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. US20040005560A1 full length cDNA

FILE REFERENCE: H1-A0106

CURRENT APPLICATION NUMBER: US/10/108,260A

CURRENT FILING DATE: 2002-03-27

NUMBER OF SEQ ID NOS: 5458

SOFTWARE: Patentlin Ver. 2.1

SEQ ID NO 1966

LENGTH: 2479

TYPE: DNA

ORGANISM: Homo sapiens

US-10-108-260A-1966

Query Match

Best Local Similarity 43.9%; Score 52.8; DB 15; Length 2479;

Matches 225; Conservative 0; Mismatches 287; Indels 0; Gaps 0;

1328 AGGACATCGACGCTGAACTTAAGAACAGCTCTCAAAACAGAACAGACGCCA 1387

1377 AGGAGAAAGAGATGTGGAGCGAGAGAGAGAGATTCACGACGAGAGATACGGAG 1436

1388 AGGTGCTTGAAGAACTCTGAAATTATGAGCGAGAACTGCTTGAATCGAGAGGATA 1447

1437 AGGAGAAAGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1496

1448 ATCGATCGTGAACAGAGAACTTAAGATGAGAGATGAACAGAGGAGGAGGAGG 1507

1497 AGGAGATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1556

1508 CACAGCAGGCTTTTTCATGATTCAATCAACAGATCCATGAAGAGAGAGCGAAAG 1567

1557 AGAAGTGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1616

1568 AGGAGAAATTCGAGATGTTGACGACGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1627

1617 AGGAGAAATGTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1676

1628 AGAATTAAATCCCTCTAGCAATGACGATTCGCGAAAGAGGAGGAGGAGGAGG 1687

1677 AGGAGAGAGAGATACGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1736

1688 TCATCGAGTTTCAAGAGAAAGAGATGAGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1747

1737 AGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1796

1748 AAGATCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1807

1797 GGGAGCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1856

1808 TGGAGAAAGAAATTTGATGAGGCTTTGGAACAG 1839

1857 TGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1888

Search completed: April 8, 2004, 20:41:55
Job time : 471.907 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 19:27:05 ; Search time 60 Seconds
(without alignments)
2943.203 Million cell updates/sec

Title: US-10-030-829-3

Sequence: 1 MSRRAGPMSEKKNVGGYRP.....EPDEALQMLYKRLHNDD 625

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_29Jan04:*

1: geneseqp1980s:*\n2: geneseqp1990s:*\n3: geneseqp2000s:*\n4: geneseqp2001s:*\n5: geneseqp2002s:*\n6: geneseqp2003as:*\n7: geneseqp2003bs:*\n8: geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3313	100.0	625	4 AAB31798	Aab31798 Amino ac
2	191	5.8	652	6 ABP98826	Abp98826 Human str
3	182	5.5	1972	2 AAB81171	Aab81171 Human BAZ
4	182	5.5	1972	6 ABR64241	Abt64241 Angiogene
5	179.5	5.4	615	6 AAE32117	Aae32117 Human cyt
6	179.5	5.4	1372	2 AAW56473	Aaw56473 Protein w
7	179	5.4	1203	4 AAM79264	Aam79264 Human pro
8	178	5.4	1203	5 ABG96296	Abg96296 Human ova
9	175	5.3	1879	4 AAM25750	Aam25750 Human pro
10	174.5	5.3	580	6 ABR57420	Abt57420 Human NOV
11	172.5	5.2	593	5 ABP62858	Abp62858 Human pol
12	172.5	5.2	933	4 ABG08600	Abg08600 Novel hum
13	172	5.2	892	7 ADB79862	Adt79862 Rat myosi
14	172	5.2	892	7 ADD47859	Add47859 Rat Prote
15	172	5.2	892	7 ADD47855	Add47855 Rat Prote
16	171.5	5.2	593	6 ABR57419	Abt57419 Human NOV
17	170.5	5.1	676	7 ADA08247	Ada08247 Human tum
18	170.5	5.1	725	2 AAW39165	Aaw39165 Human RHA
19	170.5	5.1	725	5 ABG60842	Abg60842 Human rec
20	170.5	5.1	725	5 AAU11436	Aau11436 Human hya
21	170.5	5.1	725	6 ABR57021	Abt57021 Human RHA
22	170.5	5.1	725	7 ADA08245	Ada08245 Human tum
23	170.5	5.1	725	7 ADC02450	Adc02450 Human rec
24	168.5	5.1	1192	3 AAB18165	Aab18165 Plasmodin
25	167.5	5.1	1238	4 ABB62022	Abb62022 Drosophi

26	167.5	5.1	1294	4 ABB63502	Abb63502 Drosophi
27	167.5	5.1	1898	2 AAy30795	Aay30795 A human t
28	167.5	5.1	1898	7 ADD48869	Add48869 Human Pro
29	167.5	5.1	2274	4 ABB58657	Abb58657 Drosophi
30	167.5	5.1	3899	6 ABR32048	Abt32048 Human cer
31	167.5	5.1	3917	6 ABR32050	Abt32050 Human cer
32	166.5	5.0	1979	3 AAB18171	Aab18171 Plasmodin
33	166.5	5.0	949	3 ABG15508	Abg15508 Novel hum
34	166	5.0	1162	3 AAY6255	Aay6255 Kapost's
35	166	5.0	1162	3 AAY65800	Aay65800 HIV8 ORF
36	166	5.0	1162	4 AAB62331	Aab62331 Amino ac
37	166	5.0	1162	5 ABB05621	Abb05621 Kapost's
38	165	5.0	661	6 ABP98842	Abp98842 Human str
39	165	5.0	2633	4 ABG6505	Abg6505 Novel hum
40	165	5.0	2663	4 AAM39097	Aam39097 Human pol
41	165	5.0	2688	4 AAM40883	Aam40883 Human pol
42	164	5.0	1354	2 AAM23654	Aaw23654 Physiol
43	164	5.0	1354	2 AAW71020	Aaw71020 A modifie
44	164	5.0	1354	2 AAY07082	Aay07082 Renal can
45	164	5.0	1354	6 ABU05182	Abu05182 Human exp

ALIGNMENTS

RESULT 1
ID AAB31798 standard; protein; 625 AA.
XX
AC AAB31798;
XX
DT 15-MAY-2001 (first entry)
XX
DE Amino acid sequence of the Arabidopsis SGS3 polypeptide.
XX
KM SGS3 gene; post-transcriptional inactivation; RNA degradation;
XX viral resistance; resistance; fatty acid content; protein content.
XX
OS Arabidopsis thaliana.
XX
PN WO200105951-A2.
XX
PD 25-JAN-2001.
XX
PF 13-JUL-2000; 2000MO-FR002052.
XX
PR 16-JUL-1999; 99FR-00009417.
XX
PR 26-JAN-2000; 2000FR-00001006.
XX
PA (AVET) AVENTIS CROSCIENCE SA.
XX (INRG) INST NAT RECH AGRONOMICUE.
XX
PI Beclin C, Elmayan T, Vaucheret H;
XX
DR WPI; 2001-159529/16.
XX N-PSDB; AAF25374.
XX
PT New SGS3 gene from Arabidopsis thaliana, useful for increasing virus
XX resistance in plants and, when inhibited, for increasing transgene
XX expression.
XX
PS Claim 8; Fig 1; 36pp; French.
XX
CC The present sequence represents an Arabidopsis thaliana SGS3 polypeptide.
XX The SGS3 gene is essential for post-transcriptional inactivation
XX (degradation of RNA) and for resistance to viruses. Overexpression of
XX SGS3 results in plants with increased resistance to viruses, while
XX inactivation of SGS3 in transgenic plants (e.g. by expressing antisense
XX RNA, by mutation or by homologous recombination) increases the level of
XX the transgene product. This product may e.g. impart resistance (to
XX herbicide, insects or pathogens), alter contents of essential fatty acids
XX or proteins, or is pharmaceutically active, e.g. an immunoglobulin or
XX interferon

XX Sequence 625 AA;

Query Match 100.0%; Score 3313; DB 4; Length 625;
Best Local Similarity 100.0%; Pred. No. 1.9e-250; Indels 0; Gaps 0;
Matches 625; Conservative 0; Mismatches 0;

QY 1 MSSRAGPMSEKKEKVVQGGYRPEVQLVQGLAGTGLASSQDDGGEMVYSKKKKPKGNTSG 60
DB 1 MSSRAGPMSEKKEKVVQGGYRPEVQLVQGLAGTGLASSQDDGGEMVYSKKKKPKGNTSG 60
QY 61 KTWVSONSNPPRAGWGQQGSGSNVSGRGNNSGRGNNGRGIOANISGRGALSRRYDN 120
DB 61 KTWVSONSNPPRAGWGQQGSGSNVSGRGNNSGRGNNGRGIOANISGRGALSRRYDN 120
QY 121 NFVAPPVSRPLLEGKNNQARGSAQHTVQEPFVEDVDVNDASEENDSDALDSDDD 180
DB 121 NFVAPPVSRPLLEGKNNQARGSAQHTVQEPFVEDVDVNDASEENDSDALDSDDD 180
QY 181 LASDDYDSVQSHSGSRKQNKWPKFPGSLDSLSTIEQINERQWHCPACQNGPAIDW 240
DB 181 LASDDYDSVQSHSGSRKQNKWPKFPGSLDSLSTIEQINERQWHCPACQNGPAIDW 240
QY 241 YNHPLLAHARTGARRVKLHRELAIVLEKOLMRGASVIPCETTYGQWKGLGEEKDYK 300
DB 241 YNHPLLAHARTGARRVKLHRELAIVLEKOLMRGASVIPCETTYGQWKGLGEEKDYK 300
QY 301 IYVPMVITMNTRLDKDDNDKMLGMGNQELLEYFDKYEARARHSYGGPGHGMSTLME 360
DB 301 IYVPMVITMNTRLDKDDNDKMLGMGNQELLEYFDKYEARARHSYGGPGHGMSTLME 360
QY 361 SSATGYLAEERLHRELAEMGLDRIANGQKSMFSGVROLVGLATKODLDFNQSQK 420
DB 361 SSATGYLAEERLHRELAEMGLDRIANGQKSMFSGVROLVGLATKODLDFNQSQK 420
QY 421 TRKFEKLSQGVNVKELROISDNOOLNFKNLSKQNHAVYLESLFSEKLRRTA 480
DB 421 TRKFEKLSQGVNVKELROISDNOOLNFKNLSKQNHAVYLESLFSEKLRRTA 480
QY 481 EDRIRYKORTKQHEONREEMDAHDPFMDSTIKQIHERPRAKENEEMLOQOQBARAVVQ 540
DB 481 EDRIRYKORTKQHEONREEMDAHDPFMDSTIKQIHERPRAKENEEMLOQOQBARAVVQ 540
QY 541 QOQNINPSSNDCKRAEYVSFIEFOEKEMEEFVEREMLIKQEKEMDMKKRHHET 600
DB 541 QOQNINPSSNDCKRAEYVSFIEFOEKEMEEFVEREMLIKQEKEMDMKKRHHET 600
QY 601 FDLKEPFDEALBQMLTKGHLNEDD 625
DB 601 FDLKEPFDEALBQMLTKGHLNEDD 625

RESULT 2

ABP98826 ID ABP98826 standard; protein; 652 AA.

AC ABP98826;

DT 15-JUL-2003 (first entry)

DE Human structural and cytoskeletal associated protein #17.

XX Cytosolic; antiarteriosclerotic; anticonvulsant; nootropic; antianginal;
XX neuroprotective; cerebroprotective; hypotensive; cardiact; osteopathic;
XX antiinflammatory; antiarthritic; virocid; gene therapy; human; stroke;
XX structural and cytoskeleton-associated protein; SCAP; cancer; angina;
XX atherosclerosis; epilepsy; Huntington's disease; hypertension;
XX heart failure; osteoporosis; osteoarthritis.

OS Homo sapiens.

XX MO2003031940-A2.

XX

PD 17-APR-2003.
XX
PF 10-OCT-2002; 2002MO-US032851.

PR 12-OCT-2001; 2001US-0328931P.
PR 19-OCT-2001; 2001US-0360681P.
PR 02-NOV-2001; 2001US-0343896P.
PR 09-NOV-2001; 2001US-0346308P.
PR 16-NOV-2001; 2001US-0332385P.
PR 07-DEC-2001; 2001US-0340776P.
PR 11-JAN-2002; 2002US-0347703P.

XX (INCY-) INCYTE GENOMICS INC.

PI Becha SD, Bhatia U, Blake JF, Borowsky ML, Burrill JD, Chang H,
PI Chawla NK, Elliott VS, Emerling BM, Forsythe IJ, Gorrard AE, Lal PG,
PI Griffin JA, Halalia AJA, Ho A, Ison CH, Kable AE, Knare R, Luo W,
PI Lee S, Lee EA, Lee SY, Lehr-Mason PM, Li JX, Lindquist EA, Luo W,
PI Marquis JP, Ramkumar J, Richardson TW, Sprague MW, Swarnakar A,
PI Tang YT, Warren BA, Yang J, Yue H, Zebardjian Y, Zheng W;
XX MPI; 2003-403125/38.
DR N-PSDB; ACC44313.

PT New human structural and cytoskeleton-associated proteins (SCAP) useful
PT for diagnosing, treating and preventing diseases or conditions associated
PT with the aberrant SCAP expression e.g. cancer, osteoporosis, or epilepsy.
PS Claim 1, Page 258-259; 361pp; English.

XX This sequence represents a novel isolated human structural and
XX cytoskeleton-associated protein (SCAP) polypeptide. The polypeptides and
XX polynucleotides encoding them are useful in diagnosing, treating and
XX preventing diseases or conditions associated with the decreased
XX expression or over expression of SCAP, such as cell proliferative (e.g.
XX cancer, atherosclerosis), neurological (e.g. epilepsy, Huntington's
XX disease, stroke), heart (e.g. hypertension, heart failure, angina) and
XX skeletal muscle disorders (e.g. osteoporosis, osteoarthritis) or viral
XX infections. These are also useful in assessing the effects of exogenous
XX compounds on the expression of nucleic acid and amino acid sequences of
XX SCAP. The SCAP or its fragments are useful in screening compounds for
XX effectiveness as agonist or antagonist of the polypeptides, or in
XX altering the expression of the target polynucleotide and compounds that
XX specifically bind to or modulate the activity of the polypeptide. The
XX microarray is useful in monitoring or measuring protein-protein
XX interactions, drug-target interactions, and gene expression profiles

XX Sequence 652 AA;

Query Match 5.8%; Score 191; DB 6; Length 652;
Best Local Similarity 19.0%; Pred. No. 5.8e-06;
Matches 132; Conservative 110; Mismatches 257; Indels 196; Gaps 28;

QY 74 WGGQOQGGSNVSGRGNNSGRGNNGRGIOANISGRGALSRRYDNFVAPP--VSR 130

DB 4 WKPGQRRGKEGPEARRAERGGGGG--GVAPRSPAREPRRSCLLPFWGAAMT 61

QY 131 PPL---EGW-----NWQARGSSAQHTVQEPFVEDVDVNDASE 166

DB 62 PDLNLFKKGMSILDEPEPPSPSLTTTSTSQMK-----KHWFLVLDSSLKTYRDSAE 115

QY 167 EENDSDALDSDDDLASDDYDSVQSHSGSRKQNK-----WFKKFGGS 210

DB 116 E---ADELDGDLRSCTDVEYAVQRRYGFQIHKKAVYLTLSMTSGIRRNWTEALAKT 172

QY 211 LDSLSTIEOI-----NEPQWHCPACQNGP-----GALDW 240

DB 173 VAPTAPDVTKLSQNKENALHSYSTQKGPKAGSGRASEVTSRGGRKADGQKQALDY 232

QY 241 YNLHPLL-----AHARTKARRVKLHRELAIVLENDLQNRG-----ASVTCG 283

DB 233 VELSPLTQASPORATTPARTPDRLAKQ--EELERDLAQRSEERRKWFATDSRTPEVPAG 290

19-SEP-2002; 2002MO-AU001282.
 27-SEP-2001; 2001AU-00007973.
 27-SEP-2001; 2001AU-00007974.
 11-OCT-2001; 2001AU-00008210.
 29-OCT-2001; 2001AU-00008532.
 13-NOV-2001; 2001AU-00008838.
 28-AUG-2002; 2002AU-00951032.
 (BION-) BIONOMICS LTD.
 Gamble JR, Hahn CN, Vadas MA;
 WPI; 2003-354655/33.
 N-PSDB; ACF34516.
 New angiogenic genes and polypeptides, useful for diagnosing, prognosticating or treating an angiogenesis-related disorder, e.g. cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or cardiovascular diseases.
 Claim 15; SEQ ID NO 176; 90bp; English.
 The invention relates to the isolation of novel genes (ACF34446-ACF34559) encoding proteins (ABR64180-ABR64281) involved in the process of angiogenesis. The nucleic acid molecules are useful in identifying and/or obtaining full-length human genes involved in an angiogenic process. The nucleic acid molecule, polypeptides or complexes encoded, cells or genetically modified non-human animals derived from these are useful for the screening of candidate pharmaceutical compounds used in treating, angiogenesis-related disorders. They are also useful for diagnosing, prognosticating or treating an angiogenesis-related disorder, which involves uncontrolled or enhanced angiogenesis or is a disorder in which a decreased vasculature is of benefit (e.g. cancer, rheumatoid arthritis, diabetic retinopathy, psoriasis or cardiovascular diseases such as atherosclerosis), or involves inappropriately arrested or decreased angiogenesis or is a disorder in which an expanding vasculature is of benefit (e.g. ischemic limb disease or coronary artery disease). The modulator of expression or activity of the polypeptide encoded by the nucleic acid sequence is useful for manufacturing a medicament for the treatment of an angiogenesis-related disorder. This sequence corresponds to one of the novel angiogenic protein
 Sequence 1972 AA;
 Query Match 5.5%; Score 182; DB 6; Length 1972;
 Best Local Similarity 20.8%; Pred. No. 0.00014;
 Matches 113; Conservative 83; Mismatches 190; Indels 158; Gaps 20;
 140 QARGSAGHTAVQEPFVEDVDVNASSEENDSDALDDDDDLAS-----183
 394 QFGCTGSDIPSSKSDSDNEDEDEDEDEDEDEDEDEDEDSQSSSDNSSTEGSEER 453
 184 --DDYSDVSQKSHGRKONKFKFGSL--DSISIEQINPQORWHCPACONGPAID 239
 454 DDDDKQODESDSTEGSEKTSMLKNTTSVKSPSNLTHGSTRPMLIAKAPGAPALC 513
 240 WYNIHLPL-----AHARTKAGRRVKLRLAELVLEKQ-----LQMR 275
 514 SEQSQPAFLCTSSSTLSSPHSGTSKRRRYTDEREIRILEYQWQRETRIRNFGRLQOE 573
 276 GASVIFGGEIYGW-----KGLQDEKQ-----YEIYW-- 303
 574 VAYVAPCGKTLKQVPEYIKLRSNGIMDSRDNFSSAKIRVDIFYEARDGQEQWGLL 633
 304 -----PPWVIMPTRLDKDNDKWLGNQELLEYDQY--BALPAHSYEPQGRGMS 355
 634 KEDEVIRIRAMERRRRPPNP-----DRQPARRESMRRRRKPPPVNVA 679
 356 VAMESSA--TGYIEARLRLAEMGLDIRAMQKRSMSGGVROLYGLATKQDLIDF 413
 680 EFLNADAKLRLKLOAQEIAFOAQAQIKLRLKQEQEARYVAKKQAQAIMAEF----- 734

414 NQHSQGRLLFEELK-SYQEWVVELRQISBDN-----QQLNYFNKLSKONKAKVLEES 468
 735 -----KKRQEQIIMKQOEKIRIQIQRWEKLRPAQIIEAKKKKKEBAKKLLEAE 788
 469 LIMESEKLRRTAEDNRIRYQRTKQHEQNRREMDAHRFFMDSIKQIHERDRAKEENFEM 528
 789 KEIKKEKMR-----QNAVLLKH-QERERRQH-----MLMKAMEARKKAEK--ER 833
 529 LQGEERAKVVGQOQONINPSSNDCKRAEVSSTIEFQEKMEEFYEREMLIKQDEKK 588
 834 LKQEKR-----DEKRLN-----KERRLEORLELEM-AKELKKP 866
 589 MEDM 592
 867 NEDM 870
 RESULT 5
 AAE32117
 ID AAE32117 standard; protein; 615 AA.
 AC AAE32117;
 DT 24-MAR-2003 (first entry)
 DE Human cytoskeleton-associated protein, CSAP-15.
 XX Human, cytoskeleton-associated protein, CSAP-15; atherosclerosis; cancer;
 KW gene therapy.
 OS Homo sapiens.
 FH Key Location/Qualifiers
 FT Domain 66..174
 FT /note="pH domain"
 PN WO200279404-A2.
 PD 10-OCT-2002.
 XX 25-MAR-2002; 2002MO-US009288.
 PF 29-MAR-2001; 2001US-0280508P.
 PR 03-APR-2001; 2001US-0281323P.
 PR 13-APR-2001; 2001US-0283769P.
 PR 04-MAY-2001; 2001US-0288609P.
 PR 10-MAY-2001; 2001US-0290518P.
 PR 18-MAY-2001; 2001US-0291870P.
 PR 29-MAY-2001; 2001US-0294451P.
 XX (INCY-) INCYTE GENOMICS INC.
 PA Hafalja AJA, Tang TY, Yue H, Khan FA, Ison CH, Baughn MR;
 PI Warren BA, Duggan BM, Thangavelu K, Honchell CD, Azimzai Y,
 PI Elliott VS, Burford N, Ding L, Yue H, Becha S, Emerling BM;
 PI Richardson TW, Lee SY, Bandman O, Lal PG, Lee S, Gietzen KJ;
 PI Walla NK, Griffin JA, Lee EA, Swarnakar A, Ring HZ, Jones KA;
 DR WPI; 2003-092894/08.
 DR N-PSDB; AAD49604.
 XX New human cytoskeleton-associated proteins, useful for preparing a
 PT composition for diagnosing or treating a disease or condition associated
 PT with decreased expression or overexpression of functional CSAP e.g.,
 PT cancer.
 PS Claim 1; Page 177-178; 233pp; English.
 XX The invention relates to new human cytoskeleton-associated protein (CSAP)
 CC and its polynucleotide. The polypeptide is useful for preparing a
 CC composition for diagnosing or treating a disease or condition associated
 CC with decreased expression or overexpression of functional CSAP e.g.,
 CC atherosclerosis or cancer. The present sequence is human CSAP-15 protein.

CC The invention is useful in gene therapy
XX
SQ Sequence 615 AA;

Query Match 5.4%; Score 179.5; DB 6; Length 615;
Best Local Similarity 19.0%; Pred. No. 4.2e-05;
Matches 128; Conservative 102; Mismatches 255; Indels 187; Gaps 26;

QY 74 WGGQOQGRSNNVSGRNGNGRGIGQANISGRRLSKYNNPAPP---VSR 130
DB 4 WGPGRGRKEGEPARRRAERAGGGGGG--GVAPRSPRPRPSCLLPWGAAMT 61
QY 131 PPL-EGGM-----NMQAGSGAQTAVOEPFVEDVDVNAE 166
DB 62 PDLNFKKGMSTLDEGPSPSLTTSTSQMK-----KMFVLTSLKTYRSTAE 115
QY 167 EENDSDALDDSDDLASDDYSDVQSKSHGSRKONK-----WFKKFGS 210
DB 116 E--ADELGEIDIRCTDTEYAVQRYNGFOIHTKDAVYTLTAMTSGIRRMIEALRKT 172
QY 211 LDSLSTEQI-----NRPQRMHCPACONCP-----GALDW 240
DB 173 VAPTSPADVTKLSDSKENALHSYSTQKGPLKAGEPRAGEVIRSGGPRKADGQORALDY 232
QY 241 YNLHPL----AHARKGARVYKRLRELAVELEKDLQWEG-----ASYIPCG 283
DB 233 VELSPITQASPOARKPARTPDLAKQ--BELERDLAQSEBRKMFATDSRTPEVPAG 290
QY 284 EYVGQKGLGEDEKDEYIWPVPVITMTRLDKNDKWLGMGNQELLEFDYKYLEALRA- 342
DB 291 E--GPRRGLG-----APLTEQQNRLESEIEKKQLEKLEPLRE--NKRVPETAL 336
QY 343 -----RHSYQGRHGM-----SVLMFESSATGYLEARLRELAEMGLDIANG 387
DB 337 LMQSRGERGPPSDGHEALEKEERACERSLAEMESSHQVVEELQRHNE----- 384
QY 388 QKSMFSGVROLGYFLATKODIDIFNHSQKTRKLFELKSYQEWVYKELRQISEDNOQ 447
DB 385 -----RELQRIQOEKEMV--LAETPATAATSALEAMKAQOELSRLELSK-----TRs 429
QY 448 LNYFNKRLSKONK--HAKVLEESLEIMSEKL--RTAEDNRIVRQTKMQH-----EQNREE 500
DB 430 LQQGPGLRKQHSQDVEALKRELQVLSQYSCLEIGALMROAEBREHRLRCQEGGE 489
QY 501 MDANDFPNDSTKQIHERDAKEMFEMLOQOERAVYQOQONINPSSDDCKRAEYV 560
DB 490 LLRHN-----QELHGR--LSEBIDQLRGFLASQGMGCGRSNERSCELEVLIRYK 539
QY 561 SSFIEFOEKEMEBFVEREMILKDOE-----KKMEDMKRHHSEIFDLKEFEDEA 610
DB 540 ENELOYLKKEVQCTRELQOMQDKRFTSGKYQDYVVELSHITRSEBELQLEKHLRLA 599
QY 611 LEQLMYKGLHIN 622
DB 600 MALQEKESMRN 611

RESULT 6
AAM56473
ID AAM56473 standard; protein; 1372 AA.

AC AAM56473;
XX
XX
DT 25-MAR-2003 (revised)
DT 14-AUG-1998 (first entry)

Protein with Rho protein-combining and kinase activity.

XX Rho protein-binding activity; protein kinase activity; inhibitor;
KW smooth muscle fibre formation; smooth muscle contraction;
KW circulatory disease; treatment; tumour formation; metastasis inhibitor;
KW autoimmune disease; platelet aggregation inhibitor.

OS Bos sp.
XX
PN JP10113187-A.

XX 06-MAY-1998.
XX
PF 20-NOV-1996; 97QP-00127990.
XX

PR 20-NOV-1995; 95JP-00325129.
PR 05-JUN-1996; 96JP-00017150.
PR 26-APR-1996; 96JP-00131206.
PR 24-JUL-1996; 96JP-00213243.
PR 23-AUG-1996; 96JP-00241061.
PR 20-NOV-1996; 96JP-00324594.

XX (KIRI) KIRIN BREWERY KK.

XX WPI; 1998-315475/28.
XX N-PSDB; AAV23127.

PT Bovine and human Rho protein-binding protein kinase - used to develop
PT products for treatment of smooth muscle disorders, circulatory disease,
PT cancer and autoimmune disease.

XX Disclosure; Page 35-39; 66pp; Japanese.

XX The present sequence represents a protein that has Rho protein-binding
CC activity and protein kinase activity. Inhibitors of the Rho-binding
CC protein kinase can be used to inhibit smooth muscle fibre formation and
CC smooth muscle contraction. Other applications of the Rho protein-binding
CC agent and its' products are as a circulatory disease treating agent, a
CC tumour formation or metastasis inhibitor, an autoimmune disease treating
CC agent or a platelet aggregation inhibitor (Updated on 25-MAR-2003 to
CC correct PF field.) (Updated on 25-MAR-2003 to correct PR field.)
XX

SQ Sequence 1372 AA;

Query Match 5.4%; Score 179.5; DB 2; Length 1372;
Best Local Similarity 20.8%; Pred. No. 0.00013;
Matches 116; Conservative 108; Mismatches 182; Indels 151; Gaps 27;

QY 147 QHTAVQEPFVDVDVNAEENSDSDALDDSDYSDVQSKSHGSRKONKMFKK 206
DB 509 QHNMAEYQKADHEADKKNLENDVNSLKDQJEDLK-----KKNQN----- 549
QY 207 FPGSLDSLSEIENEPQRMHCPACONGGALDWTNLHPLAHARTKAR----- 256
DB 550 -----SQISTEKVNOLORO-----LDETV--ALRTESPTRARLRKTOAESK 590
QY 257 -----RVKLRHRE--LAVLEKDLQMR--GASVIPGGEIYQGW 289
DB 591 QIQOLESNNRDIQDNKCLLETAFLKLEKFINQSVLESERDRTHGSEIT--NDIQGKI 648
QY 290 KGLGEDEKDEYIWPVPVITM-----TRLKDDNDKVIANG-----NOBLEFFDYK 337
DB 649 SGLIEDVKNKGLILAKVLEKRGQLOERFTDLKEKKNKMIIDWYQKVIQQLSEGETEH 708
QY 338 EALRARHSYGPQGRHGMSTVLMFESSATGYLEARL--HRELAEMGLDIANGOKR-SNF 393
DB 709 KATKARLADKNKIYESIE--EAKSEAKMEKELSEBRTLKQKVENLLEEKSCSIL 764
QY 394 SSGVROLGYFLATKODL-----DIFNHSQKTRKLFELKSYQEWVYKELRQISEDNO 446
DB 765 DCDLKO-----SQQINELKQKDVNLNDVRLT-LKIQEOTQKCLTV--NDLMKQTV 815
QY 447 QLTNYFK--NKLKONKHAQVLEESLEIMSEKLRTAEDNRIVRQTKMQHQRNREMDA 503
DB 816 QVNTLMKSEKQKQENHLLKMKMSLEKQNAELRKE-----RQADQGMKELQQLBA 868
QY 504 HRRF--FNDSTKQIHERDAKEMFEMLOQO-----ERAKVYQOQONINPSSDD-C 553
DB 869 EGYFSTLYKTQVRELKECEERTKLCLELOQKQKQELQDERSILAQLITTLTKADESOLA 928

QY 554 RRAEVSFFIEPQEKMEFVEERMLIKDQK-----KMDKKRHHHEIFDL- 603
DB 929 RSIABQYSDLE-KEKIMKE-LEIKEMARHKOELTEKDATIASLETNRILTSDVANLA 986
QY 604 -EKE-----FDEALEQL 614
DB 987 NEEKELINKKEHQBOL 1003

RESULT 7
AAM79264

ID AAM79264 standard; protein; 1203 AA.

AC AAM79264;
XX 06-NOV-2001 (first entry)
XX Human protein SEQ ID NO 1926.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KM tissue growth factor; immunomodulatory; cancer; leukaemia;
KM nervous system disorder; arthritis; inflammation.

XX Homo sapiens.
XX MO200157190-A2.
XX 09-AUG-2001.
XX 05-FEB-2001; 2001MO-US004098.

XX 03-FEB-2000; 2000US-00496814.
XX 27-APR-2000; 2000US-00560875.
XX 20-JUN-2000; 2000US-00598075.
XX 19-JUL-2000; 2000US-00620325.
XX 01-SEP-2000; 2000US-00654936.
XX 15-SEP-2000; 2000US-00663561.
XX 20-OCT-2000; 2000US-00693325.
XX 30-NOV-2000; 2000US-00728422.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT, Asundi V, Zhou P, Xu C, Cao Y;
PI Ma Y, Zhao QA, Wang D, Wang J, Zhang J, Ren F, Chen R, Wang ZW;
PI Xue AB, Yang Y, Wejman T, Goodrich R;

XX WPI; 2001-476283/51.
XX N-PSDB; AAK52397.

XX Nucleic acids encoding polypeptides with cytokine-like activities, useful
PT in diagnosis and gene therapy.

XX Claim 20; Page 4327-4329; 6221pp; English.

XX The invention relates to polynucleotides (AAK51456-AAK53435) and the
CC encoded polypeptides (AAM78333-AAK80302) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: Records for SEQ ID NO 2110 (AAK52581), 2111
CC (AAK52582) and 3666 (AAM80020) are omitted as the relevant pages from the
CC sequence listing were missing at the time of publication

XX Sequence 1203 AA;

Query Match 5.4%; Score 179; DB 4; Length 1203;
Best Local Similarity 19.0%; Pred. No. 0.00012;

Matches 173; Conservative 121; Mismatches 284; Indels 332; Gaps 40;

QY 15 QGGRPE-----VEQLVGLAGTR--LASSQDDGGEWEVSKKNKPKNGTSKTY 63
DB 37 RGGRRPADARASTYVAVRVQAGAPFVVLNSGEGGDSFGVQITGANDOG--ASGA-- 93
QY 64 VSQNSNPRAWGCGOQGRGNSVSGRGNVSGRGNNGRGI-----QANISGRGRLSRKYD 119
DB 94 LSSDLLEPENYSQVKFPAPSGSTSDPEBGAAYNGKILRSHSQASLAGGVPDPNSRS 153
QY 120 NNFV-APPVSRP-----PLE-----GGMWQARGSGAAGHTAVGEFPDVE 158
DB 154 NSMLETAPKVASPGSTIDTAPLSSVDSLINFDSQLG--QARGKGRRTM--LPBQ 208
QY 159 DDVDVASBEENDSDALDSDDD-----LASPDYSDVSQKSHGRKQKWFKEFGSLDS 213
DB 209 RRGSKSLDRSLRPTDFEEREROSTMHTSTSTKYDNHTGTSKQPAQSN-----LSP 259
QY 214 LS-----IEQINERQOMHCPACON-----GPGALDWY--NLH 244
DB 260 LSGFRSRQTDQWVLQSFEBPRRSADPTMLQFKSTPDLRDQEAAPGSDVDMKATY 319
QY 245 PLLAARTKGAARVYLHRELALEVELEKIDQMRGASVIFG--LYGQ-----WKG 291
DB 320 GILREGSSESTSVR--RKVSIVLEK--WQPLVWVSGSTKAVAGGELTRKYBELORK 374
QY 292 LGEDEKDYELVWPVVIIMNTRFLDKDNDKWLGMGNQELLE-----Y 333
DB 375 LDEYVKRQKLEPQVGLERQLEKTEBECNRL-----QELLEKRGKGAQSNKELQNMRL 430
QY 334 FDKYEALR-----ARHSYGFQ----- 349
DB 431 LQGEDLRHGLETVQWELQNKLVHQGEPFAKELLKDLLETRELLEVELEGQVVEBOL 490
QY 350 --GHGMSVL--MESSAT-----GYLEAR-- 371
DB 491 RLREBELTALKALKEEVASRDQVEHYRQYQRTBQLRSMQATODHAIVLEBRQM 550
QY 372 -----LHRELAEMGLDRIAMGOKRSMFSGVRQLYGLATKODL----- 410
DB 551 SALVRLGLQRELEETBEETGRW--QSMFOKNKEDL--RATKQBLQLRMEKEEMEELIG 604
QY 411 --DIFNQH-----SQGKTR--LKPELASYGMVYEL----- 438
DB 605 EKIEVLQRELEQARASAGTRQVEVYLKKELLRTQE--ELKEIQAEHQSGDEVAGRHRDELE 663
QY 439 -----ROISEDNQO-----LNVFKKLSKONGKAKVLEESLEIMSEK----- 475
DB 664 KQALAVLRVEADRGRELEBQNLQKTLQQLRODCBESAKKAVLAELATVLDGRRAAVET 723
QY 476 -LRTAEDNRIVRQRTKQMHQNRREM-----DAHDFEMDSIKQIHERRDAKEBNEM 528
DB 724 TLRETEQENDEPFRRLTGLEQQLKETRGVLVDGEAVEARLRLKRLAEKQLEBALNA 783
QY 529 LQOQE-----RAKV-----VGQOQONINPSSNDCRKAEEVSFFIEFOE- 568
DB 784 SQREBSLAAAKRALPARLEAQRGARIGQOQTLNRLLEEGRQREVLTRGKAELEQ 843
QY 569 -----KEMEFAVEEREMLIKDEKEMDKKHHHEIFLEKEF--DEALEQLM 615
DB 844 KRLIDTVDRLNKELEKIGEDSKQALQQLQALQEDVKEKARREVDAGQQAQDMWSEAK 903
QY 616 YKHGLHNEED 625
DB 904 TSGGLRLQD 913

RESULT 8
ID ABG96296 standard; protein; 1203 AA.

XX ABG96296;
XX

11-DEC-2002 (first entry)
Human ovarian cancer marker OV9.
Human, ovarian cancer; marker; cancer; familial history; brain disorder; central nervous system disorder; bacterial meningitis; viral meningitis; Alzheimer's disease; Parkinson's disease; cerebellar oedema; hydrocephalus; brain herniation, inflammation, encephalitis, testicular disorder; non-tuberculous granulomatous orchitis; connective tissue disorder; heart disorder; ischaemic heart disease; atherosclerosis; neoplasm; histological type; carcinogenic; ovarian cancer marker.
Homo sapiens.
WO200271928-A2.
19-SEP-2002.
14-MAR-2002; 2002WO-US007826.
14-MAR-2001; 2001US-0276025P.
14-MAR-2001; 2001US-0276026P.
10-AUG-2001; 2001US-0311732P.
19-SEP-2001; 2001US-0323580P.
26-SEP-2001; 2001US-0324967P.
26-SEP-2001; 2001US-0325102P.
26-SEP-2001; 2001US-0325149P.
(MILL-) MILLENNIUM PHARM INC.
Morahan JE, Gannavaram M, Hoerch S, Kamathkar S, Kovatis SG; Meyers RE, Morrissett MP, Olandt PJ, Sen A, Vieby PO, Mills GB, Baer RC, Lu K, Schmandt RE, Zhao X, Glatt K; WPI; 2002-723277/78.
N-PSDB; ABS763388.
Assessing whether a patient is afflicted with ovarian cancer, useful in assessing the stage or progression of the disease, comprises comparing the expression level of a cancer marker in a sample from a patient and from a non cancer patient.
Disclosure; Page 160-163; 481pp; English.
The present invention relates to a new method for assessing whether a patient is afflicted with ovarian cancer. The method involves comparing the expression level of a marker in a patient sample and the normal level of expression of the marker in a control non-ovarian cancer sample, where the marker is selected from 363 cancer markers described in the specification. The method of the invention is useful in diagnosing or characterising cancer, in detecting the presence of cancer as early as possible, and the recurrence of ovarian cancer. The method may also be of particular use with patients having an enhanced risk of developing ovarian cancer (e.g. patients having a familial history of ovarian cancer). The cancer markers may be used in the management and treatment of e.g. brain and central nervous system disorders (e.g. bacterial and viral meningitis, Alzheimer's disease or Parkinson's disease), brain disorders (e.g. cerebral oedema, hydrocephalus or brain herniations), inflammations (e.g. bacterial or viral meningitis or encephalitis), testicular disorders (e.g. non-tuberculous granulomatous orchitis), connective tissue disorders, or heart disorders (e.g. ischaemic heart disease or atherosclerosis). The compositions and methods may also be used in assessing the histological type of neoplasm associated with ovarian cancer, monitoring the progression of ovarian cancer, determining whether ovarian cancer has metastasized or is likely to metastasize, selecting a composition for inhibiting ovarian cancer, assessing the ovarian carcinogenic potential of a compound, or inhibiting ovarian cancer or at risk of developing ovarian cancer. The present amino acid sequence represents one of the ovarian cancer markers described in the invention
Sequence 1203 AA;

Query Match	5.4%;	Score 178;	DB 5;	Length 1203;		
Best Local Similarity	19.0%;	Pred. No. 0.00014;				
Matches	173;	Conservative	121;	Mismatches	264;	
			Indels	332;	Gaps	40;

QY	15	QGVYRP-----VEQLVQGLAGTR--LASSQDDGGEWEVISKKNKPKRGTSKWTW	63
DB	37	RGGRNPAMDARASTGYVAVRVQGIAGCFVVLNIGKEGKDSFGVQIKGANOQ--ASGA--	93
QY	64	VSQNSNPRAWGGQQOQGRGNSVSGRGNVSGRGNNGRGIT-----QNNISGGGRALSRKYD	119
DB	94	LSSLDLELPENFYGVGKFPAPSGQSTDEBPAAWNGILKSHSGVASLAGEPPVDPNSRS	153
QY	120	NNFY-APPVYRP-----PLE-----GGMNQARQSAQHTAVQCFPQVE	158
DB	154	NSMELAPKVASPQSTIDTAPLSSVDSLINFQDQGG--QAKRGTRRTRM--LPFEQ	208
QY	159	DDVNNAGEEENSDALDDSDDD-----LASDDYSDVSQKSHGRKKQMKPFKFGSLDS	213
DB	209	RKRSKSLDSRLPRTLPFERERQSTNTNMTSTSKYDNHVGTSKQPAQSQ-----LSP	259
QY	214	LS-----TEQINPQRQMHCPACN-----GPGALDW--NLH	244
DB	260	LSGFSRSRQTDWVLQSGFEEPRRRAQDPMTLQFKSTPDLRLQDEAARPGSDVHMKATY	319
QY	245	PLNHARTKARRKYLRLAEVLKQLDMGASVIRPGE---LYGQ-----WKG	291
DB	320	GILREGSESETSYR--RKSLVLEK--MPLVWSSGSTKAVAGQSELTTRYVELOK	374
QY	292	LGDEKDYEIYMPVVIIMTRLRKDDNDKMYGMQNSLLE-----Y	333
DB	375	LDEEVKKRQKLEPSQVGLEROLEETECBSRL--QELLRKGEAQAOSNKELQNMRL	430
QY	334	FDKYEARL-----ARSHYGPQ-----	349
DB	431	LDQGEDLRHGLQVWELQNKTKVQGPAPKAVLLKDLTRELLEVLBEQKQVEQL	490
QY	350	--GHRGMSVL--MFESSAT-----GLVEER--	371
DB	491	RLRRRELTALMGALKEEYASRDQVEYHVRQOYQDTEQLRRSMODATQDHAVLEARQK	550
QY	372	--LHRELAENGLDRIAMQKRSMPSSGAVROLYGFATKQDP-----	410
DB	551	SALVRGLQRELEETSEETGHW--QSMFQKNKEDL--RATKQELLQRMKEKEBEBELG	604
QY	411	--DIFNOH-----SQGKTR--LKEELSYQEMVVKEL-----	438
DB	605	EKIEVLQRELEQARASAGDTRQVEVLKKELLRTQE--ELKEIQARQSQEVAGRHNRLE	663
QY	439	--ROISEDNOQ---LNFYKQKLSKQNHAKVLEESLEINSEK-----	475
DB	664	KQLAVLRVEADRGRELBEONLQLOKTLQQLRQDCEEAQKAVVAEAEATVLAGORAAVET	723
QY	476	--LRTAEENRIVRQRTKQHEONDEEM-----DAHNDPMDSTIKQIHERRDAKEENEM	528
DB	724	TLRELTQEBENDEFRRRIIGLEQOLKETRGLVDGGEVAVELRLDPKQORLEAKQOLEBALNA	783
QY	529	LQOOB-----RAKY-----VGOOQONINSSNDCKRAEYVSFTFQES	568
DB	784	SQEBEGSLAAKRALEARLEBAQGLARLAGQEQOTLNALBEEGQQRVLRGRKALEBEQ	843
QY	569	--KEMEEFVEEREMILIKQEKMEDMKRGHHEIFDLKEF--DEALBQLM	615
DB	844	KRLLDRTVRLNKELEKIGEDSKQALQQLQALQLEDYKEXKARREVAADARQADWASAEAK	903
QY	616	YKHGLHNEDD	625
DB	904	TSGGLSRLQD	913

RESULT 9	
AAAM25750	
ID	AAAM25750 standard; protein; 1879 AA.
XX	

AAIM25750. 16-OCT-2001 (first entry)
Human protein sequence SEQ ID NO:1265.
Human; cancer; ulcer; HIV infection; human immunodeficiency virus; antiinflammatory; antithematic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; vitruide; anti-HIV; fungicide; antimutagen; cardiovascular; antiaemic; anaemia; antiagregant; haemostatic; vlnarary; antilcer; osteopathic; eczema; dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; infection; immunostimulant; gene therapy; antisense therapy; vaccine; inflammation; antiaphylactic; rheumatoid arthritis; septic shock; pancreatitis; cardiac dysfunction; neuropathology; cardiac anaphylaxis; autoimmune; genetic disease; haematopoietic disorder; platelet disorder; asthma; thrombocytopenia; osteoporosis; severe combined immunodeficiency; allergic rhinitis; diabetes; multiple sclerosis; depression; Alzheimer's disease; Parkinson's disease; neurodegenerative disorder; neurological disorder.
Homo sapiens.
WO200153455-A2.
26-JUL-2001.
22-DEC-2000; 2000WO-US035017.
23-DEC-1999; 99US-00471275.
21-JAN-2000; 2000US-0048725.
25-APR-2000; 2000US-00552317.
(HXSE-) HXSEQ INC.
Tang YT, Liu C, Dermanac RT;
WPI; 2001-457603/49.
N-PSDB; AAM99691.
Isolated human polynucleotides encoding polypeptides, useful for the treatment and diagnosis of e.g. cancer, ulcers and HIV infection.
Claim 20; Page 262; 1217p; English.
AAM99166 to AAM99904 encode the human proteins given in AAM25225 to AAM25963. The proteins can have activities based on the tissues and cells they are expressed in, such as: antiinflammatory; antithematic; antiarthritic; immunosuppressive; antibacterial; endocrine; cardiant; central nervous system; vitruide; anti-HIV; fungicide; antimutagen; cardiovascular; antiaemic; antiagregant; haemostatic; vlnarary; antilcer; osteopathic; dermatological; antiallergic; antiaesthetic; antidiabetic; cytostatic; neuroprotective; antidepressant; nootropic; antiparkinsonian; and immunostimulant. The proteins and polynucleotides encoding them can be used in gene therapy, antisense therapy and vaccine production. The proteins and polynucleotides are useful for screening for agonists or antagonists of a protein and for the treatment and diagnosis of disorders associated with the activity of a protein e.g. inflammation, rheumatoid arthritis, septic shock, pancreatitis, cardiac dysfunction, neuropathology, cardiac anaphylaxis, viral, bacterial, HIV and fungal infections, autoimmune, genetic diseases, haematopoietic disorders, anaemia, platelet disorders, thrombocytopenia, wounds, burns, ulcers, osteoporosis, severe combined immunodeficiency, eczema, allergic rhinitis, asthma, diabetes, cancer, multiple sclerosis, depression, Alzheimer's disease, Parkinson's disease, neurodegenerative and neurological disorders
Sequence 1879 AA;
Query Match 5.3%; Score 175; DB 4; Length 1879;
Best Local Similarity 20.2%; Pred. No. 0.00046;
Matches 116; Conservative 94; Mismatches 194; Indels 170; Gaps 24

QY	158	EDVDVNASEF-----	ENDSDALDDSDDDLASDDYDSDVSQKSHSRKQNKPKFK	206
Db	1297	QKEIENITQQEERKAAYDKLEKTKTKRLQOELDLVLV-DLDNQRQLVSNLEKKQRK-FDQ		1354
QY	207	PFPSIDLSISTEQINPEPQWCHPCACQNGPGLIDWYMLHPLLAHRTGADRVKLRLELAE	266	
Db	1355	LLAEKKNTSSKTVADERDR-----	VEAEKREKETKALSLARLAE	1393
QY	267	VLEKDLQMRASVYIPCGEI-----	YGQWKGLGED-EKDYEIYMPWVIMNTRLD	315
Db	1394	ALEAKSELENTNMKLKAKWGRPSGASKQDVGQ--ELSHDLEKSKALGDPRLEEMKTQLE		1451
QY	316	KDDNDKMLGNGOEL-----	FDYKELARHSHYGQGHKGMV	356
Db	1452	E-----	IGRTIELASPRRDATKRLRELVNNQAPSRASFER-DLQARTQEONESRRHLQR	1501
QY	357	IMPFSSATGYLAEARHLRELAEMGLDIRAWGQKRSM-----	FSGG--VRQLYGFLA	405
Db	1502	QLHEYTEE--LEDERKQRLALAAAKITLGMDPVPTITLDLXADSALKGCGKALIKQLRKLOA		1559
QY	406	TKODL-----	DIF--NQHSGKTRLFELKSYQEWVVKELRQISEDNQOLNY	450
Db	1560	QMKDFQELBDARASRDEIFATAKENKKAKSLIADIMQLOE--DLAAAEGRKQADL		1615
QY	451	PKYKLS-----	KQNKAKV--LEESL-----EIMSEKLR-----	478
Db	1616	EKEELAEELASSLSGRNALODEKRRLELRARLAQLEELLEEEOGNENASDRKRAKTTQQAEO		1675
QY	479	-----	TAEDNRIIVPORTKMGQONEEMDAHDFPMDISIKQIHERDAR-----EE	524
Db	1676	LSNELATERSTQGNESARQOLEQNKELSKLHEMGAVKSKSTIALLAKITAOLEE		1735
QY	525	NFEMLQOQEPRAKVVGQOQONINPSSNDCKRAAEVVSFLIEFQKMEBEPEEENB----		579
Db	1736	QVEOEAREKQOATKSLKQKD-----	KKKEILLQVEDERKMAEQYKQEAQKGNAR	1785
QY	580	-MLIKDEKKMKEDMKKRNHEEIPDLKEPFEALE	612	
Db	1786	VKQLKQLEEAHEESQRINANRRKLQRELDPEATE		1819
RESULT 10				
ABR57420				
ID	ABR57420	standard; protein; 580 AA.		
XX	ABR57420;			
AC				
XX				
XX	15-SEP-2003	(first entry)		
DT				
XX				
DE	Human NOV4b protein SEQ ID NO:18.			
XX				
XX	Human; NOV4; cytosstatic; cardiant; antiinflammatory; immunosuppressive;			
KW	antiallergic; haemostatic; anti-HIV; antidiabetic; antilatterioleleotic;			
KW	anorectic; antisthmatic; nephroretoric; antiarthritic; hepatotropic;			
KW	neuroprotective; nootropic; antibacterial; vincicde; antiparasitic;			
KW	relaxant; anticonvulsant; hypotensive; vasodiloric; antiparkinsonian;			
KW	vulnerary; angiogenic; antiangiogenic; gene therapy; vaccine; cancer;			
KW	cardiomyopathy; atherosclerosis; hypertension; diabetes; inflammation;			
KW	autoimmune disorder; allergy; blood disorder; AIDS; obesity; asthma;			
KW	acquired immunodeficiency syndrome; nephropathy; cirrhosis; arthritis;			
KW	Alzheimer's disease; Parkinson's disease; goitre; infection; stroke;			
KW	muscular dystrophy; epilepsy; wasting disorder; chromosome 5.			
OS	Homo sapiens.			
XX				
XX	WO200294870-A2.			
XX				
PD	28-NOV-2002.			
XX				
PF	02-NOV-2001; 2001WO-US051580.			
XX				
PR	02-NOV-2000; 2000US-0245291P.			

CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 593 AA;

Query Match	5.2%;	Score 172.5;	DB 5;	Length 593;
Best Local Similarity	19.0%;	Pred. No. 0.00014;		
Matches 110;	Conservative 100;	Mismatches 209;	Indels 161;	Gaps 23

```

OY      162 DNASEENDSDALDDSDDDLSDVDSVSQKSHSRKONK-----WK 205
Db      52 DSTAEE---ABEIDSEIDLRCSTDTVEAVQANRYFQJHTKDVAVYL$AMTSGIRNMIE 108
OY      206 KFGSSLDLSLIEOI-----NEBQROMCBACONGP-----235
Db      109 ALRKTVPRTSAADVTYKLSDSNKENALH$YSTCKGLPKAGEORAGSEVISRGPRKADGQR 168
OY      236 GAIDWYNLHPLL-----AHARTGARVRKYLRBLAEVLEKDLQMFQ-----AS 278
Db      169 QALDVELSPITLOASPQRARFTPARFPDRILAKO--BELEDLRLQREBERKKWEATDSTRTP 226
OY      279 VIPCEIYGOMXGLGEDEKDYEIWPMPWIIMNTRILDKDNDKM-----322
Db      227 EYPAGE--GPRRGLG-----APLTEDQNHLSEIEIKKMOELEKPLRENKRVPVL 274
OY      323 LGMGNO-----ELLEYFDKYEALP$AR-HSYGPOGRHMSVLMFE$SA---TG 365
Db      275 TALNLQ$RGERGRCPSPDGHEALE--KEVALPAQLAEWLRLQ$BAPASALRSOEODGHIPBG 332
OY      366 YLEARLHRBELAENG-----LDR$ANGOKREMFSGGV$QOLYGLFLATRKDDLI$FNHG$QG 419
Db      333 YISOACECSLAE$MESSHOOWMEELORHNHERELQ$R-----LOOEKEMILAETTAAT 383
OY      420 KTRLKFELKSVOEWMYKTELROI$EDNQOLNYFANKL$SKONK--HAKYLBESLEIMS$EKL-R 477
Db      384 ASATAMKKAUYEBEL$REL$K---TRSLQOGPDGLRKQHGSV$EALMKRELYLS$OYQ 439
OY      478 RTAEDNRIRVORTKMOH-----EONKEENDANDHFRFMDSIKOIHERRDAKENEFMLQOO 532
Db      440 KLELGALMRQ$BERHTLRRCOQEGOCOLLRNH-----QELHGR---LSEIEDOLRGF 489
OY      533 BEAKVVGOQOQININPSSNDCKRRAEVSV$FIEFOEKEME$FUEEREMLIKODE-----586
Db      490 IASOGGNCGNCGR$NERS$CELEVTLRVKNEIEOLYKKEVAC$LFDELQOMOXDRFTS$GY 549
OY      587 -----KNMEDMKKHNEEIFDL$EKPDEAL$EOLMYUKGILNH 622
Db      550 QDVUYELSHIKTRSERIEOLKHL$MALMAALOEXSMRN 569

```

RESULT 12	
ABG08600	
ID	ABG08600 standard; protein; 923 AA.
XX	
AC	ABG08600;
XX	
DT	13-FEB-2002 (first entry)
XX	
DE	Novel human diagnostic protein #8591.
XX	
KW	Human; chromosome mapping; gene therapy; forensic;
KW	food supplement; medical imaging; diagnostic; genetic disorder.
XX	
OS	Homo sapiens.
XX	
PN	WO200175067-A2.
XX	
PD	11-OCT-2001.
XX	
PF	30-MAR-2001; 2001WO-US008631.
XX	
PR	31-MAR-2000; 2000US-00540217.
PR	23-AUG-2000; 2000US-00649167.
PR	

XX (HYSE-) HYSEQ INC.
PA
XX
XX Drmanac RT, Liu C, Tang YT,
PI
XX
XX WPI; 2001-639362/73.
DR N-PSDB; AAS72787.
DR
XX

PT New isolated polynucleotide and encoded polypeptides, useful in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits and to assess biodiversity.

PS Claim 20; SEQ ID NO 38959; 103pp; English

The invention relates to isolated polynucleotide (I) and polypeptide (II) sequences. (I) is useful as hybridisation probe, polymerase chain reaction (PCR) primers, oligomers, and for chromosome and gene mapping, and in recombinant production of (II). The polynucleotides are also used in diagnostics as expressed sequence tags for identifying expressed genes. (II) is useful in gene therapy techniques to restore normal activity of (II) or to treat disease states involving (II). (II) is useful for generating antibodies against it, detecting or quantitating a polypeptide in tissue, as molecular weight markers and as a food supplement. (II) and its binding partners are useful in medical imaging of sites expressing (II). (I) and (II) are useful for treating disorders involving aberrant protein expression or biological activity. The polypeptide and polynucleotide sequences have applications in diagnostics, forensics, gene mapping, identification of mutations responsible for genetic disorders or other traits to assess biodiversity and to produce other types of data and products dependent on DNA and amino acid sequences. ABG00010-ABG03177 represent novel human diagnostic amino acid sequences of the invention. Note: The sequence data for this patent did not appear in the printed specification, but was obtained in electronic format directly from WIPO at http://wipo.int/pub/published_pct_sequences

SQ Sequence 923 AA:

Query Match	5.2%	Score 172.5;	DB 4;	Length 923;
Best Local Similarity	19.7%;	Pred: No. 0.00027;		
Matches 60; Conservative	64;	Mismatches 112;	Indels 69;	Gaps 9

```

QY      368 EAERLRELAEMGLDIRIANGCQRSMSPGSGVRLQYGLATKCOLDIJFNOSCKTILKFE- 426
D8      153 ETNHLKRELESVG-----RLOAEVENVNMOLSLNROEERLREQEER 195
QY      427 -----LKSYOENVVKELRQISEDNQOLNFKNLSKONKHAYVL 465
D8      196 LKEQERQOEQEDRLHEQEERLREQEERLCEQEERLREHEERLCEQEERLREQ 255
QY      466 EESLEIMSEKRLRTAE-----DNRLVRQSTKQOHEQNR-----EENDADHRF 508
D8      256 EERLCEQEERLREHEERLCEQEERLCEQEERLHEQEERLREQEERLREHEERL 315
QY      509 MDSIKOIHERRDAKENFEMLPQOEAKYVVGQOQONINSSNDCKR-----AEVY 560
D8      316 CEQEERLREHEERLCEQEERLREQEER--LCEQEERLREQEERLCEQEERLREQ 373
QY      561 SSTRLEQEK-EMERFVBEREMLIKQOEKQME--DMKRHHEELFIDLEKEFDEALEQOLMYKH 618
D8      374 EKLLEERRQOEQERLLERRLLEVEKLLLEQROOEERHEMLRERLEE-VEKLLLEQ 432
QY      619 GLANNE 623
D8      433 RRHEE 437

```

RESULT 13
ADB79862
ID ADB79862 standard; protein; 892 AA
XX
AC ADB79862;

XX 04-DEC-2003 (first entry)
 XX Rat myosin heavy chain, SEQ ID 102.
 DE Analgesic; pain; streptozocin-induced diabetes; rat.
 XX Rattus norvegicus.
 XX EP1279744-A2.
 XX 29-JAN-2003.
 XX 26-JUL-2002; 2002EP-00255249.
 XX 27-JUL-2001; 2001GB-00018354.
 XX 07-FEB-2002; 2002GB-00002910.
 XX (WARN) WARNER LAMBERT CO.
 XX Brooksbank RA, Dixon AK, Lee K, Pinnock RD,
 XX WPI; 2003-395407/38.
 XX N-PSDB; ADB79863.
 XX Use of isolated gene sequences and encoded polypeptides that are
 PT upregulated in the spinal cord in response to streptozocin-induced
 PT diabetes for screening compounds for the treatment of pain, or for
 PT diagnosing pain.
 XX Claim 1; Page 181-183; 334pp; English.
 XX The present invention relates to nucleotide sequences which are useful in
 CC the screening of compounds for the treatment of pain, or for the
 CC diagnosis of pain. The nucleotide sequences are up-regulated in the
 CC spinal cord in response to streptozocin-induced diabetes. The present
 CC sequence was used to illustrate the invention.
 XX Sequence 892 AA:
 SQ
 Query Match 5.2%; Score 172; DB 7; Length 892;
 Best Local Similarity 21.0%; Pred. No. 0.00028;
 Matches 122; Conservative 98; Mismatches 186; Indels 174; Gaps 29;
 QY 141 ARGGSAQHTRAVQFPP-DVEDVDNMASEENDSDALDSDS--DDDLASDDYSDVSQKSHGS 197
 DB 259 SRKQAPFTQOIEELKRLQEEVAKSAL--AAHALQSSRHDCDLREQYESEAKAEIQ 315
 QY 198 RKONK-----WPKFFGSLDSLSTEQINEPQRMHCPACQNGPAIDWYNLHPLLAHA 250
 DB 316 RANSKANSEVAQWRRTKY-----ETDAIQRTSE-----LEEA 346
 QY 251 RTYGARVKLHRELAEVLEKDLQWRGASVIPCSEIYGQWKGLGEDEKDYEIWPPWYIM 310
 DB 347 KKTLAQRDLDAEBEHAIVN-----AKCASLEKTRQLQNEVEDL-----MIDVE 390
 QY 311 NNR-----LDKDNNDKMLGNGNELLEYFKYKALAHSHYGGQHRGKSVLMF----- 359
 DB 391 RTVAACALDKQRN-----FDKTLLEWQKQYETTHAELEASQKESRSSTELFKIKNAV 445
 QY 360 ----ESSATGYLAEERLHELAEMLDRIAWQGR-----SMFGGVAVQLGFL-ATKQDL 410
 DB 446 EBSLDQLFTIKRNNKLLQGEISPL-TGQIAEGSKRIHELEKIKQIQEISELQALAEER 504
 QY 411 DIFNQSQGK-TLKLELSYQEMVAVKEQLQISEDNOQLYFKNKLSKONKAKVLEEST 469
 DB 505 EASLEHEEGKILRIQLEINQVKEISDRKIAEKDEIDQL-----KRN-HIRVYESMQ 555
 QY 470 EIMSELRRAENRIRVQRTKMGHQHONREMDAH--DRFMSIK-----QI 515
 DB 556 STLDAAIR--SRNDAIRIKKRGEGDNEWELQJNHSNRPAALNRYNRQTGILKDTOL 612
 QY 516 H-----ERR-----DAK-EENFEMLQOQERAKVVGQOO----- 542

DB 613 HLDLALRGQEDLKEQLAMVERRANILQAEIEELRATLEQTERSRSKIAEQELDASERVOL 672
 QY 543 ---QINPSSNDDCRRAAEVSSFFIFQEKMEEFVE-----REWL 581
 DB 673 LHTQN---TSLINTFKRRLTETDISQI---QGEEMDIVQEARNAEKKAKAITDAMMAEEL 726
 QY 582 IKDOEKK--MEDWKKRHHIEIFDLKEKFEPALEQLMYKKG 619
 DB 727 KKEQDTSALIERKKKRLBQTVKDLQHRLEDA-EQLALKGG 765
 RESULT 14
 ADD47859
 ID ADD47859 standard; protein; 892 AA.
 XX ADD47859;
 XX 29-JAN-2004 (first entry)
 DT Rat Protein AAB29713, SEQ ID NO 13555.
 XX Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 KW chronic constriction injury; CCI; spared nerve injury; SNI; Chung.
 XX Rattus norvegicus.
 XX WO2003016475-A2.
 XX 27-FEB-2003.
 XX 14-AUG-2002; 2002WO-US025765.
 XX 14-AUG-2001; 2001US-0312147P.
 PR 01-NOV-2001; 2001US-0346382P.
 PR 26-NOV-2001; 2001US-033347P.
 XX (GEHO) GEN HOSPITAL CORP.
 PA (PARB) BAYER AG.
 PA Woolf C, D'urso D, Belfort K, Costigan M,
 PI WPI; 2003-268312/26.
 DR GENBANK; AAB29713.
 XX New composition comprising two or more isolated polypeptides, useful for
 PT preparing a medicament for treating pain in an animal.
 XX Claim 1; Page; 1017pp; English.
 XX The invention discloses a composition comprising two or more isolated rat
 CC or human polynucleotides or a polynucleotide which represents a fragment,
 CC derivative or allelic variation of the nucleic acid sequence. Also
 CC claimed are a vector comprising the novel polynucleotide, a host cell
 CC comprising the vector, a method for identifying a nucleotide sequence
 CC which is differentially regulated in an animal subjected to pain and a
 CC kit to perform the method, an array, a method for identifying an agent
 CC that increases or decreases the expression of the polynucleotide sequence
 CC that is differentially expressed in neuronal tissue of a first animal
 CC subjected to pain, a method for identifying a compound which regulates
 CC the expression of a polynucleotide sequence which is differentially
 CC expressed in an animal subjected to pain, a method for identifying a
 CC compound that regulates the activity of one or more of the
 CC polynucleotides, a method for producing a pharmaceutical composition, a
 CC method for identifying a compound or small molecule that regulates the
 CC activity in an animal of one or more of the polypeptides given in the
 CC specification, a method for identifying a compound useful in treating
 CC pain and a pharmaceutical composition comprising the one or more
 CC polypeptides or their antibodies. The polynucleotide or the compound that
 CC modulates its activity is useful for preparing a medicament for treating
 CC pain (e.g. spinal segmental nerve injury (Chung), chronic constriction
 CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene
 CC therapy). The sequence presented is a rat protein (shown in Table 2 of

CC the specification) which is differentially expressed during pain. Note:
 CC The sequence data for this patent did not form part of the printed
 CC specification, but was obtained in electronic form directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 892 AA;

Query Match 5.2%; Score 172; DB 7; Length 892;

Best Local Similarity 21.0%; Pred. No. 0.00028; Indels 174; Gaps 29;
 Matches 122; Conservative 98; Mismatches 186;

141 ARGSSAQTAVOFP-DVEDVDVNAEENDSDALDD--DDDLASDDYSDVSGKSHGS 197
 259 SRGKQAFVTOQIEHLKQLEBEVAKSAL--AHALQSSRHDCDLREOYEEOEAKALEQ 315
 198 RKONK-----WFKRFGSLDSLSEIQINEPORQWHCPACONGPAILDWNHPLLAHA 250
 316 RAMSKANSEVAQWRTKY-----ETDAIQRTBE-----LEBA 346
 251 RTKGARRVKLHRELAEVLEKDLQWRGASVIPCSEIYQMGKLGDEKDYEIWPPVITM 310
 347 KKCLAQRLQDAEHEVAVN-----AKCASLEKTKQRLQNEVEDL-----MIDVE 390
 311 NTR-----LDKDDNDKMLGNGNOELLEYPDKYELARHSHYSGQGRHSGSVLMF----- 359
 391 RTNACALDCKQKN-----FDKILAEWKQCYETTHALEASQKESKSLSTELFKINAY 445
 360 -----ESSATGYLEAERLHRELAEVLEKDLQWRGASVIPCSEIYQMGKLGDEKDYEIWPPVITM 410
 446 EESLDQLETLKREKNLQOEISDL--TEQIABGGRIHELEKIKQIOEKSSELQALALEA 504
 411 DIFNHSQOK--TRLFELKSYQEWNVKELROISEDNOQLYFNKLSKONKHAQVLEESL 469
 505 EASLEHSGKILRIQLEINQVSEIDRKIAEKDEBIDOL-----KRN-HIRVESMQ 555
 470 EIMSKRLRTADNRIVQRTMGHONREEMDAH--DRFMDISIK-----QI 515
 556 STLDIAEIR--SRNDAIRIKKKMEGDLNEMELQUNHSHRMAAEALRYNRNQGILKDTOL 612
 516 H-----ERR-----DAK-EENFEMLOOQERAKVVGQO----- 542
 613 HLDLDRGDELDKEQLAWERRANLLQAEIELRATLQTSRSRKIAEOELLDASERVOL 672
 543 ---QINPSSNDCKRRAEVSFIEPQEKHEEVEE-----REML 581
 673 LHTQN---TSLINTKKKQLETDISQI--QGEWEDIVQEARNAEERAKKATITDAAMABEL 726
 582 IKDOEKK--MEDMKKXHHHEIFDLKEKPEDEALQOLMYKHG 619
 727 KKEQDTSALHERKKQLBQTVADLQHRUDEA--EQALAKG 765

RESULT 15

ADD47855

ADD47855 standard; protein; 892 AA.

ADD47855;

29-JAN-2004 (first entry)

Rat Protein AAB29713, SEQ ID NO 13551.

Rat; pain; neuronal tissue; gene therapy; spinal segmental nerve injury;
 chronic constriction injury; CCI; spared nerve injury; SNi; Chung.

Rattus norvegicus.

MO2003016475-A2.

27-FEB-2003.

14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P;
 PR 01-NOV-2001; 2001US-0346382P;
 PR 26-NOV-2001; 2001US-033347P.

XX (GENO) GEN HOSPITAL CORP.
 PA (FARB) BAYER AG.

XX Woolf C, D'Urso D, Befort K, Costigan M;

XX WPI; 2003-266312/26.
 DR GENBANK; AAB29713.

XX New composition comprising two or more isolated polypeptides, useful for
 preparing a medicament for treating pain in an animal.

PS Claim 1; Page: 1017pp; English.

The invention discloses a composition comprising two or more isolated rat
 or human polynucleotides or a polynucleotide which represents a fragment,
 derivative or allelic variation of the nucleic acid sequence. Also
 claimed are a vector comprising the novel polynucleotide, a host cell
 comprising the vector, a method for identifying a nucleotide sequence
 which is differentially regulated in an animal subjected to pain and a
 kit to perform the method, an array, a method for identifying an agent
 that increases or decreases the expression of the polynucleotide sequence
 that is differentially expressed in neuronal tissue of a first animal
 subjected to pain, a method for identifying a compound which regulates
 the expression of a polynucleotide sequence which is differentially
 expressed in an animal subjected to pain, a method for identifying a
 compound that regulates the activity of one or more of the
 polynucleotides, a method for producing a pharmaceutical composition, a
 method for identifying a compound or small molecule that regulates the
 activity in an animal of one or more of the polypeptides given in the
 specification, a method for identifying a compound useful in treating
 pain and a pharmaceutical composition comprising the one or more
 polypeptides or their antibodies. The polynucleotide or the compound that
 modulates its activity is useful for preparing a medicament for treating
 pain (e.g. spinal segmental nerve injury (SNi)) in an animal (e.g. gene
 therapy). The sequence presented is a rat protein (shown in Table 2 of
 the specification) which is differentially expressed during pain. Note:
 the sequence data for this patent did not form part of the printed
 specification, but was obtained in electronic form directly from WIPO at
 ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 892 AA;

Query Match 5.2%; Score 172; DB 7; Length 892;

Best Local Similarity 21.0%; Pred. No. 0.00028; Indels 174; Gaps 29;
 Matches 122; Conservative 98; Mismatches 186;

141 ARGSSAQTAVOFP-DVEDVDVNAEENDSDALDD--DDDLASDDYSDVSGKSHGS 197
 259 SRGKQAFVTOQIEHLKQLEBEVAKSAL--AHALQSSRHDCDLREOYEEOEAKALEQ 315
 198 RKONK-----WFKRFGSLDSLSEIQINEPORQWHCPACONGPAILDWNHPLLAHA 250
 316 RAMSKANSEVAQWRTKY-----ETDAIQRTBE-----LEBA 346
 251 RTKGARRVKLHRELAEVLEKDLQWRGASVIPCSEIYQMGKLGDEKDYEIWPPVITM 310
 347 KKCLAQRLQDAEHEVAVN-----AKCASLEKTKQRLQNEVEDL-----MIDVE 390
 311 NTR-----LDKDDNDKMLGNGNOELLEYPDKYELARHSHYSGQGRHSGSVLMF----- 359
 391 RTNACALDCKQKN-----FDKILAEWKQCYETTHALEASQKESKSLSTELFKINAY 445
 360 -----ESSATGYLEAERLHRELAEVLEKDLQWRGASVIPCSEIYQMGKLGDEKDYEIWPPVITM 410
 446 EESLDQLETLKREKNLQOEISDL--TEQIABGGRIHELEKIKQIOEKSSELQALALEA 504
 411 DIFNHSQOK--TRLFELKSYQEWNVKELROISEDNOQLYFNKLSKONKHAQVLEESL 469

```

Db      505 EASLEHEEGILRIQELNOVKSEIDRKIAEKDEIDOL-----KRN-HIRVESMQ 555
Qy      470 EIMSEKLRRTAEDNRIVRORTKQOHEONREEMDAH--DRPFMSIK-----OI 515
Db      556 STDAEIR---SRNDAIRIKKKMEGDLNEMEIQLNHSNRMAAEALRNYRNTOGILKDTOL 612
Qy      516 H-----ERR---DAK-BENFEMLOQOERAKVVGQOQ-----542
Db      613 HLDALRGQEDLKEQOLAMVERRANLQAEIEBLRATLEQTERSRKIAEQELDASERVOL 672
Qy      543 ---QINPSNDCKRAEVSFIEFOEKEMEEFVE-----REML 581
Db      673 LHTON---TSLINTYKKLETDISOI---QEMEDIYQEARNAEBEKAKKAITDAAMAEEL 726
Qy      582 IKDOEKK--MEDMKGRHHEEIFDLEKEFDEALBOLWYKHG 619
Db      727 KKEODTSAHLERMKNLQETVKDLOHRLDEA-BQLALKGG 765

```

Search completed: April 6, 2004, 19:33:25
 Job time : 65 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 6, 2004, 19:31:20 ; Search time 21 Seconds
(without alignments)
2862.843 Million cell updates/sec

Title: US-10-030-829-3

Perfect score: 3313

Sequence: 1 MSRRAGPMKEKNVGGYRP.....EPDEALQMLKKHGLNEDD 625

Scoring table: BLOSUM62

Gapop 10.0 , Gapept 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Maximum Match 0%

Listing first 45 summaries

Database :
1: PIR 78:*
2: PIR1:*
3: PIR2:*
4: PIR3:*
5: PIR4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	233.5	7.0	644	2	T46211 hypothetical prote
2	211	6.4	634	2	E96840 hypothetical prote
3	192	5.8	634	2	E86293 T24D18.1 protein -
4	186.5	5.6	554	2	E85015 hypothetical prote
5	183	5.5	559	2	T01724 hypothetical prote
6	180.5	5.4	885	2	T03446 probable transcrip
7	180	5.4	662	2	T01533 hypothetical prote
8	179.5	5.4	1388	2	S70633 serine/threonine-s
9	178.5	5.4	1957	2	A59294 skeletal myosin -
10	174	5.3	1937	2	I38055 myosin heavy chain
11	173.5	5.2	978	2	A70387 conserved hypotet
12	173	5.2	488	2	T21596 hypothetical prote
13	173	5.2	676	2	S00084 myosin heavy chain
14	172.5	5.2	1156	2	B70356 chromosome assembl
15	172.5	5.2	1407	1	S28589 trichobialin - rab
16	172.5	5.2	1871	2	D96796 probable heat choc
17	171.5	5.2	2116	2	A26655 myosin heavy chain
18	171	5.2	1432	2	B85431 trichobialin like
19	171	5.2	1948	2	S00485 gene 11.1 protein
20	170.5	5.1	725	1	UC5016 hyaluronan recepto
21	169	5.1	1938	2	A59293 skeletal myosin he
22	168.5	5.1	1192	2	A71623 probable secreted
23	168.5	5.1	1957	2	A45627 myosin heavy chain
24	167.5	5.1	1898	1	A45973 trichobialin - hum
25	167	5.0	1085	2	F96712 hypothetical prote
26	166.5	5.0	1979	2	C71622 hypothetical prote
27	166	5.0	736	2	D86271 protein F16A14.2 l
28	166	5.0	1088	2	T18559 hypothetical prote
29	165	5.0	1313	2	A48467 myosin heavy chain

30	165	5.0	1931	2	A59234 slow myosin heavy
31	165	5.0	2663	1	S28261 centromere protein
32	164	5.0	1354	2	S69211 serine/threonine-s
33	164	5.0	1790	2	S67593 transport protein
34	163.5	4.9	980	2	E71606 hypothetical prote
35	163	4.9	1426	2	A59287 hypothetical prote
36	163	4.9	1940	1	A59287 myosin heavy chain
37	162	4.9	1992	2	S02771 myosin heavy chain
38	161.5	4.9	1738	2	T14867 interapkin - slime
39	161.5	4.9	1875	2	S38173 myosin-like protei
40	161	4.9	451	2	G70241 hypothetical prote
41	161	4.9	936	2	S39083 myosin heavy chain
42	161	4.9	1039	2	S18199 myosin heavy chain
43	161	4.9	1837	2	T41023 probable nuclear p
44	160.5	4.8	621	2	S10450 myosin heavy chain
45	160.5	4.8	1642	2	T08880 NMDA receptor-bind

ALIGNMENTS

RESULT 1

T46211
hypothetical protein T8P19.180 - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 04-Feb-2000 #sequence_revision 04-Feb-2000 #ext_change 04-Feb-2000
C/Accession: T46211
R/Choiene, N.; Robert, C.; Broctier, P.; Wincker, P.; Catclicco, L.; Artiguenave, F.; Sai
Submitted to the Protein Sequence Database, December 1999
A/Reference number: Z23008
A/Accession: T46211
A/Status: Preliminary
A/Molecule type: DNA
A/Residues: 1644 <CHO>
A/Cross-references: EMBL:AL13315
A/Experimental source: cultivar Columbia; BAC clone T8P19
C/Genetics:
A/Map position: 3
A/Intons: 133/1; 373/3; 403/3; 496/3; 566/3
A/Note: T8P19.180

Query Match 7.0%; Score 233.5; DB 2; Length 644;
Best Local Similarity 22.9%; Pred. No. 2.8e-05;
Matches 110; Conservative 88; Mismatches 192; Indels 91; Gaps 19;

QY	181	LASDDYDSVQKSHSKRKNKFKFPGSLDSLSIQINEPQRQNHCPACQNGPACIDW	240
DB	7	LSDDSDSDSDISES-----EMDEYGDWKYINLKGKGLKVLRLSPQ-AFTCPYCPNKKTSFQ	60
QY	241	YNHPLLAHARTG-----ARRVKLHRELAEVLEKDIQMRGASVTPCGEITYGQWKL	292
DB	61	YK--DLLQHSVGVSNSNDKRSKASHALATKYVQDDASASEPSSKKQKGNP1	118
QY	293	GDEKDEYIWPMPVIMTRLDKNDKWLGNQBEL-EYFDK-YEALRASHSYGPOG	350
DB	119	QDCDHEKLVYPRKGVINPTTKADGRSAGSSGKLREIYLRGNPFRVRYWYLG	178
QY	351	HRGMSVLMPESSATGYLAEERLHRELAEMGLDIRIAMQKRSWFGVRYQLYGLATKYDL	410
DB	179	HSSTAIVFEMKDNWNGHLNGLFPKAYTVDGHGKDWLKK---DGPRLGYGWIARADY	234
QY	411	---DIFNHSQGRTRLPK-----ELKSVOEMVVKELROISEDNQQLNFKNLSKONKH	461
DB	235	NGNNITIGENIRKGTGDTTALTEEDARKOELVONLRQVLE-----KKKD	281
QY	462	AKVLESLEIMSEKLRRTAEDNRIVRQTMQHEONREEMDAHREPMDSIKQIHERRDA	521
DB	282	MKEIEELCSYKSEL-----NQLMEKKNQOKHYR-ELNATQERTSMHIQIVD----	330
QY	522	KEENFEMLOQOQERAKV-----GQOQONINSSNDCKRAEVSFIEFOE	568
DB	331	DHEKRLKLESEERKLEIKCNELAKREVNHGTERMTL---SBDLONASKNS-DELA	385

Oy		569	KEMEFVFERMFLIDQCKKMEDMKKRHHLEFDEKEPFE-----ALEOL-----MYK	617
			:: : :: :	:
			:: : :: :	:
			:: : :: :	:
			:: : :: :	:
Db		386	MEOGKADEVVKKLAEQNDEL-----HEKTIRLERORDKOALIEVEQLKGGLNVMK	438
Oy		618 H	618	
Db		439 H	439	

RESULT 2
E96940
hypothetical protein F23A5.14 [imported] - Arabidopsis thaliana
C/Species: Arabidopsis thaliana (mouse-ear cress)
C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C/Acession: E96940
R/Theologias: A.; Eckert, J.R.; Palm, C.J.; Federici, N.A.; Kaul, S.; White, O.; Alonso
Chn, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.
ansen, N.F.; Hughes, B.; Hulzar, L.

Native 408, 816-820, 2000

A:/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C
C./I., J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luoro, J.S.; Maiti, R.; Marziali
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:/Authors: Salzbeger, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A./Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A./Reference number: AB6141; PMID:21016719; PMD:11130712

A/Acession: E96940

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-634 <STO>

A/Cross-references: GB:AEO05173; NID:g6503291; PIDN:AAF14667..1; GS/PDB:GN00141

C/Genetics:
A/Gene: F23A5.14
.Map position: 1

[illegible]

DB 417 EIQELGKGLKWKH--EDED 435

RESULT 3

E86293

T24D18.1 protein - Arabidopsis thaliana

CSpecies: Arabidopsis thaliana (mouse-ear cress)

CDate: 02-Mar-2001 #sequence_revision 02-Mar-2001 #ext_change 09-Nov-2001

CAccession: E86293

RTheologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Jensen, N.F.; Hughes, B.; Hultzer, L.

Nature 408, 816-820, 2000

AAuthors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.; Cizzo, L.; J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matti, R.; Marzilli, R.R.; Rooney, T.; Rowley, D.; Sakano, H.

AAuthors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I. Ker, M.; Wu, D.; Yu, G.; Frazer, C.M.; Venter, J.C.; Davis, R.W.

ATitle: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

AReference number: A86141; MUID:21016719; PMID:11130712

AAccession: E86293

AStatus: preliminary

A:Molecule type: DNA

A:Residues: 1-634 <STO-

A:Cross-References: GB:AE005172, NID:g6587797, PIN:AAPI8488.1, GSPDB:GN00141

C:Genetics:

A:Map position: 1

Query Match 5.8%; Score 192; DB 2; Length 634;

Best Local Similarity 22.1%; Pred. NO. 0.0046;

Matches 98; Conservative 80; Mismatches 179; Indels 86; Gaps 20;

QY 225 QMHCACONGPDAIWMYNLHPLLAHA-----RTGARRVKKHLRELAETLEKDLQGRGA 277
 Db 38 QLRCPFCA-GKKKQD-YKKKELYAHTATGVSXKSATSLAQKANHLLAMLENNEL-AGYA 94
 QY 278 SVIPGGEIYGQWKLGEDEKDEYIWPVPVILIMNTRLDXDNDKMLGMNGQELLEYFDKY 337
 Db 95 EPPRPAPVPPLQDLETEPPIPHNVVYVPMMGIVNPL--KEADDKELLIDSAWMLQTLISKF 152
 QY 338 EALRRARHSYCPQCHGMSVLMFESSATGYLEARRHLRELAEMQLDIRAMQCKRSMFSG-G 396
 Db 153 KPIEVNAFMVEDDSIVGVIAKTNQDMSGFGATELKEKEFTQSSKKEMTER---SGDS 208
 QY 397 VROLGYFLATKQDL-----DIFNQSQGKTRLKFELKSYQE--MVVKELROI---SE 443
 Db 209 ESKATGWCARADDPESQPIGELYLSKEGQLRTVSDISQKNVQDRNRYLELSMIAINTNE 268
 QY 444 DNOQLNFRNKLKLS-----KONKRAVLSBSLEIMSBEKRLRLADNRIVRORTMOH 494
 Db 269 DLNKQVSYNRRYAMSLQRLVLEDKKHLQAFADETKKIQOMSLRHI--QKILYDEKELSN 325
 QY 495 EQNREEMDAHDPFMDSTIKQIHERNDAKENPEMLQQQRAKVGQGGQQQINSSNDDCR 554
 Db 326 ELDKQRDLBSH---AKQL-----EKHALYELDRQKI-----DEDK 359
 QY 555 KRAEVSFFIEFOEKMEFEVEREMILIKQEKMDKKRHHIEFIDLEKEFD--EAL 612
 Db 360 RKSDAMNSSLQASNEQKKADESVLRVVEHQRQKEDALNK----ILLLEKQDITKQTL 415
 QY 613 -----QIMYHGHLEHDD 625
 Db 416 MEIQELKGLQVM-KH-LGDDDD 436

RESULT 4
 EBS015
 hypotheetical protein AT4G01180 [imported] - Arabidopsis thaliana
 CSpecies: Arabidopsis thaliana (mouse-ear cress)
 CDate: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 16-Feb-2001
 CAccession: EBS015
 R:Anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring
 Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant *Arabidopsis thaliana*.
 A:Reference number: AB5001; MID:20083488; PMID:10617198
 A:Accession: E85015
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-554 <STO>
 A:Cross-references: GB:NC_001268; NID:g7267615; PIDN:CAB80927.1; GSPDB:GN00140
 C:Genetics:
 A:Gene: AT4G01180
 A:Map position: 4

Query Match 5.6%; Score 186.5; DB 2; Length 554;
 Best Local Similarity 23.5%; Pred. No. 0.0078;
 Matches 83; Conservative 73; Mismatches 154; Indels 43; Gaps 16;

QY 270 KDLQMGASVIFGCELYGCKGIGEDSKYETIYPMVIMTRLDKDNDRKLGKNGOE 329
 DB 2 KKKQKQEFVAETSEKNNNNECGDQOK-RYWPWVGIVANVPTEVEPSGRVGRSGST 60

QY 330 LLEYF--DKYEALRAHSYGPOGHRGMSVLMFESSATGYLEARLHRELAEMGLDRIAMG 387
 DB 61 LKDEFLTKGFNPFRVAKPIWMTKHTGFALVEFAKDKGFESAMQFEK---SPDLDRHG-- 115

QY 388 OKRSMFSG-GVR--QYGFATKQDLDIFNQSQK-TRLKFEKSYQEMVYKELQISE 443
 DB 116 -KRWKKGRLRDLKLYGMLAREDDYN--RSDTVGKGVKKKRDLSIQVEEDQKLYH 172

QY 444 --DNQOLNFYKXKLSKONKHAQYLE--ESLEIMSEKLRPAEDNRIVRQRTKQHEONRE 499
 DB 173 LFENMCQTIEKKQKQKQOEKQVDETLSELEFNMLNLSYOE--IQCKEKKMQEFTQO 230

QY 500 EMDAHRPFMDSIKQIHERDAKEENFEMLQOERAKVQGOQONINPSSNDDCRKRAE 559
 DB 231 VLQGEHKSFAE-----LEAKREKLD---ERARLI-EQALIKNEEMETRLIREM 276

QY 560 VSSFIEFOKEMEBFVEEREMLIKDOEKMDKXKHHEIFDLKEFDEALE 612
 DB 277 I-----QKMWCEONAEAMKLAEKIQKE-KEKLRKIMEMKAKNETQOE 321

RESULT 5
 T01724
 hypothetical protein A_I0002N01.10 - *Arabidopsis thaliana*
 C:Species: *Arabidopsis thaliana* (mouse-ear cress)
 C:Date: 19-Feb-1999 #sequence_revision 19-Feb-1999 #text_change 22-Oct-1999
 C:Accession: T01724
 R:Schneet, P.; Maggii, L.
 submitted to the EMBL Data Library, June 1997
 A:Description: The sequence of A. thaliana IG002N01.
 A:Reference number: Z14407
 A:Accession: T01724
 A:Status: translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-559 <SCH>
 A:Cross-references: EMBL:AF007269; NID:g2191126; PID:g2191132; ATSP:A_I0002N01.10
 C:Genetics:
 A:Gene: ATSP:A_I0002N01.10
 A:Intons: 122/1; 216/1; 276/3; 304/3; 398/3; 464/3

Query Match 5.5%; Score 183; DB 2; Length 559;
 Best Local Similarity 22.9%; Pred. No. 0.012;
 Matches 85; Conservative 72; Mismatches 140; Indels 74; Gaps 17;

QY 270 KDLQMGASVIFGCELYGCKGIGEDSKYETIYPMVIMTRLDKDNDRKLGKNGOE 329
 DB 2 KKKQKQEFVAETSEKNNNNECGDQOK-RYWPWVGIVANVPTEVEPSGRVGRSGST 60

QY 330 LLEYF--DKYEALRAHSYGPOGHRGMSVLMFESSATGYLEARLHRELAEMGLDRIAMG 387
 DB 61 LKDEFLTKGFNPFRVAKPIWMTKHTGFALVEFAKDKGFESAMQFEK---SPDLDRHG-- 115

QY 388 OKRSMFSG-GVR--QYGFATKQDLDIFNQSQK-TRLKFEKSYQEMVYKELQISE 443

DB 116 -KRWKKGRLRDLKLYGMLAREDDYN--RSDTVGKGVKKKRDLSIQVEEDQKLYH 172
 QY 444 --DNQOLNFYKXKLSKONKHAQYLE--ESLEIMSEKLRPAEDNRIVRQRTKQHEONRE 499
 DB 173 LFENMCQTIEKKQKQKQOEKQVDETLSELEFNMLNLSYOE--IQCKEKKMQEFTQO 230

QY 500 EMDAHRPFMDSIKQIHERDAKEENFEMLQOERAKVQGOQONINPSSNDDCRKRAE 559
 DB 231 VLQGEHKSFAE-----LEAKREKLD---ERARLI-EQALIKNEEMETRLIREM 276

QY 560 VSSFIEFOKEMEBFVEEREMLIK--DOEKMDM-----KKRHHEEF 601
 DB 263 -----NEEEMKRLREMIQKMWCEONAEAMKLAEKQASTSLKEKELKRLIM 315

QY 602 DLKEFDEALE 612
 DB 316 EWEAKNETQOE 326

RESULT 6
 T03446
 probable transcription regulator protein - sorghum
 C:Species: *Sorghum bicolor* (sorghum)
 C:Date: 24-Mar-1999 #sequence_revision 24-Mar-1999 #text_change 29-Oct-1999
 C:Accession: T03446
 R:Chen, M.; SamMiguel, P.; Bennetzen, J.L.
 Genetics 148, 435-443, 1998

A:Title: Sequence organization and conservation of SH2/A1-homologous regions of sorghum
 A:Reference number: Z14952; MID:98133900; PMID:9475753
 A:Accession: T03446
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-895 <CHE>
 A:Cross-references: EMBL:AF010283; NID:g2735839; PIDN:AB94013.1; PID:g2735841
 C:Genetics:
 A:Intons: 271/1; 311/3; 356/3; 454/1; 509/3; 571/3; 588/3; 642/3; 696/3; 724/2; 787/2

Query Match 5.4%; Score 180.5; DB 2; Length 895;
 Best Local Similarity 21.5%; Pred. No. 0.029;
 Matches 85; Conservative 79; Mismatches 154; Indels 77; Gaps 17;

QY 241 YMLPPLAAR-----TKGARRVKLHRELAELYEKDQMGAS--VIFGCELYGQMGK 291
 DB 27 YKSSSLVQASGVSAPNKAKEKAHRLAFYKLKNDLAKSSBPQPLVIFV-----EP 79

QY 292 LGEDEKDYETIYPMVIMTRLDKDNDRKLGKNGOELLEYFDKYEALRAHSYGPOG 351
 DB 80 QPLQNRDEKFWPMWGIIVNPTEWKDG-ROIGESGNRLKEQLSHPCPLKVIPLMTFRDH 138

QY 352 RQMSVLMFESSATGYLEARLHRELAEMGLDRIAMQKSMFSGVROLYGFATKQDLD 411
 DB 139 TGNALVEFGKDNNGFNNAATFESHFAGGFGKDKWTGKNQGS-----ELYGLAARD-- 192

QY 412 IFNQSQK---TRLKFEKSYQEMVYKELQISEDNQOLNFYKXKLSKONKHAQYLE 467
 DB 193 ---YNSPGIILADYLRKNGDLKSVNDLAKGAR---KTDRLVANLANQIEVKORYLELE- 245

QY 468 SLEIMSEKLRPAEDNRIVRQRTKQHEONREEMDAHRPFMDSIKQIHERDAK--EEN 525
 DB 246 -----SKYSETTASLEKMGKQREBOLQSYNEE-----ISKQQLARRHSQKYLIDN 291

QY 526 FEMLQOERAKVQGOQONINPSSND-DCR-KRAEVSSFIEFOKEMEBFVEEREMLIK 583
 DB 292 -QKLSSELEAKM-----NDLDVRSQDLBLAKSDYDRNLQO--EKQNAIK 336

QY 584 DOEKKMDKXKHHEIFDLKEFDEALEQOLMYKH 618
 DB 337 SSKLKATYLE-----QKADENVLKLEKH 361

RESULT 7
 T01533
 hypothetical protein A_I0005110.22 - *Arabidopsis thaliana*

A:Experimental source: sex female
C:Superfamily: myosin heavy chain; myosin motor domain homology
F:87-773/Domain: myosin motor domain homology <MMO>

Query Match 5.4%; Score 178.5; DB 2; Length 1957;
Best Local Similarity 22.1%; Pred. No. 0.095;
Matches 120; Conservative 91; Mismatches 206; Indels 125; Gaps 24;

157 VEDDVNASEENDSDALDSDSDVSDVSOQSHGRKQNMFKKFPGLSLSLSI 216
1019 VEDDKGNYSKKLKGK--LEQSLDDL-----EDNLEKRRNEIKQKRIQSEL-KVAQ 1070
217 EQINEPQRMWCPACONGPGAIIDWYMLHPPLAHARTKGAARV-----KLR----- 262
1071 ENNEBIEERQNH-----EIE-SNLKKKETEQAQITALEEEDPLVGLTKKQVNETON 1120
263 ---ELAEVLEKDIOMGASVIFGGEIYGKKGGE--DEK-----DYELVMPWVITIMN 311
1121 RITELEEEBENERQSRKAKASDLORELEBEGDRLEDEGAGTAAQVEVNRKREAEALAK 1180
312 TRLD-----KDNNDKMLGWNQELLEYFPKYALNARHSYGFQGRGMS 355
1181 LRDLDEBANNHNNQALIRKKNDAVAELGDD--IEQVQKAKAKIEKDMQAQREKQDL 1238
356 VLMFESSATGYLBAERLHRE-LAEMGLDRIAWGQKSMFGVYQULYGFATK-----QD 409
1239 VAIIDDET-----AENNNKKLAKQYEMQIA--ELQTKCDEQNRQLOEFTTLKTRLNSEN 1291
410 LDIENHQSCK-----TRLKFEIKSYQEMVVKELQISEDNQOLNFKKSLKQNH-- 461
1292 TDLGKQIEEASFQVNMATRLKAQLTSQLBEARRSLDEARD-----RNLLAQMQNYQ 1344
462 -----AKYLESELEINSEKLRRTAEENRIVRQ-RTKQNH--QNEEMADNRFPFMDSI 512
1345 HEIEQVLESHEEIEEGISELMKQDISRNABEIQQWQTFESEGILKGELESKKRQNHKM 1404
513 KQIHERRDAKEENFEMLOOQERAKVWGQOQONINPSSNDCCRRAREEVSFFIEQEKEME 572
1405 NEIQETLIDANSKISLSLEKSKRLV-----SDLDQAQMDVRANVYAQOLEKQK 1454
573 EYVEEREMLIKQEKKMEDN-----KKRHHEIFDLKEPDEALQMYKGLHN 622
1455 GF-----DKVIDEKKTKTDIATEVDNAQREARVSTELFKLKEQDEVELETTI---EGLRR 1507
623 ED 624
1508 EN 1509

RESULT 10
138055
myosin heavy chain, perinatal skeletal muscle - human
N:Contains: myosin ATPase (EC 3.6.4.1)
C:Species: Homo sapiens (man)
C:Date: 17-May-1996 #sequence revision 17-May-1996 #text change 19-Apr-2002
C:Accession: I38055; S12459; S09332; A30220; S49478
R:Nullian, E.H.; Kelly, A.M.; Pompidou, A.U.; Hoffman, R.; Schifflino, S.; Stedman, H.H.
Eur. J. Biochem. 230, 1001-1006, 1995
A>Title: Characterization of a human perinatal myosin heavy-chain transcript.
A:Reference number: I38055; MUID:95324556; PMID:7601129
A:Accession: I38055
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-1937 <RES>
A:Cross-references: EMBL:238133; NID:9558666; PID:CAA86293.1; PID:9558666
R:Karsch-Nisch, I.; Peghali, R.; Shows, T.B.; Leitwand, L.A.
Gene 89, 289-294, 1990
A>Title: Generation of a full-length human perinatal myosin heavy-chain-encoding cDNA.
A:Reference number: JH0154; MUID:90323631; PMID:2373371
A:Accession: JH0154
A:Molecule type: mRNA
A:Residues: 1-14, 'A', 16-859 <KAR>
A:Cross-references: GB:Y00821

A:Experimental source: skeletal muscle
R:Bober, E.
Submitted to the EMBL Data Library, January 1989
A:Reference number: S12458

A:Accession: S12459
A:Molecule type: mRNA
A:Residues: 502-1071, 'N', 1073-1250, 'DGG', 1253-1376, 'NT', 1379-1913, 'D', 1915-1937 <BOB>
A:Cross-references: EMBL:X51592; NID:929465; PID:CAA35941.1; PID:929466
A:Experimental source: clone gTMC-F
R:Bober, E.; Buchberger-Seidl, A.; Braun, T.; Singh, S.; Goedde, H.W.; Arnold, H.H.
Eur. J. Biochem. 189, 55-65, 1990
A>Title: Identification of three developmentally controlled isoforms of human myosin heavy chain.
A:Reference number: S09331; MUID:90235862; PMID:1691980
A:Accession: S09332
A:Molecule type: mRNA
A:Residues: 502-547, 'X', 549-617, 'X', 619-687, 'X', 689-757, 'X', 759-827, 'X', 829-897, 'X', 899-917, 'NT', 1379-1386, 'X', 1388-1456, 'X', 1458-1526, 'X', 1528-1596, 'X', 1598-1666, 'X', 1668-1736
A:Cross-references: EMBL:X51592
R:Peghali, R.; Leitwand, L.A.
J. Cell Biol. 108, 1791-1797, 1989
A>Title: Molecular genetic characterization of a developmentally regulated human perinatal myosin heavy chain.
A:Reference number: A30220; MUID:89234168; PMID:2715179
A:Accession: A30220
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 860-969, 'Q', 971-1246, 'H', 1248-1260, 'G', 1262-1296, 'Q', 1298-1503, 'AH', 1506-1846
A:Cross-references: GB:Y00821; NID:934663; PID:CAA68757.1; PID:934664
C:Genetics:
A:Gene: GDB:MYH8
A:Cross-references: GDB:125267; OMIM:160741
A:Map position: 17pter-17p12
C:Superfamily: myosin heavy chain; myosin motor domain homology
C:Keywords: actin binding; ATP; coiled coil; hydrophobic; muscle contraction; nucleotide binding; 181-188/Region: nucleotide-binding motif A (P-loop)
F:51-588/Region: actin binding #status predicted
F:658-680/Region: actin binding #status predicted
F:842-1282/Region: S2 #status predicted
F:698,708/Active site: Cys #status predicted

Query Match 5.3%; Score 174; DB 2; Length 1937;
Best Local Similarity 19.3%; Pred. No. 0.16;
Matches 121; Conservative 99; Mismatches 201; Indels 206; Gaps 26;

104 QANISGRALARKYNNVAPRPVRRPPLLEGMMQAGSAGAHNAVQEPVEDVDVN 163
1280 RARLQTEAGEYSRQDEKDALVSQSR-----SKQASTQIEELKRLQLEB 1324
164 ASEEND-SDALDSDS--DDLASDDYSDVSOQSHGRKQNMFKKFPGLSLSLSIQTN 220
1325 ETKAKYALAHALQSSRHDCDLREQYEEBOEKAEIQR-----ALSKAN 1368
221 EPQRQNHCPACONGPGAIIDWYMLHPPLAHARTKGAARVYLHRELALEKDIOMGASVY 280
1369 SEVAQV-----RTKYETDAIQTELEBAKKLAQR----- 1399
281 PCGEIYGKKGGEDEPKQYEWPMVITIMN--RLDKNDNKMVG----- 326
1400 -----IQAEHEVAANAKCASLEKTKQRIQNEVEDLMDVERSNAACAALDKK 1448
327 ---NOELLEYPDKYALNARHSYGFQGRGMSVLMF-----ESSATGYLBAERL 372
1449 QNRFDKVLSQWKQYETQAELEASQKERSISTELFKXNVYBESLDQLETRRNKKYL 1508
1509 QOEISDL-TEQIAEGGKQ-----IHELEKIKKQYQEKCEIQALAEFAASLEHEGKI 1561
421 TRLKEFLKSYQEMVVKELQISEDNQOLNFKKSLKQNHAKYLESELEINSEKLRRA 480
1562 LRIQLELNQVKSVDPKAKKEIDQL-----KRN-HRVVETMOSTLDAETR--- 1609
481 EDNRIVRQRTKQNHQNEEMDAH--DRPFMDSIK-----QTH----- 516

QY 524 ENFEMLOOQERARVAGOOQINPSSNDCKRAAEVSSFFIEQEKMEEFVEREMLIK 563
 DB 406 SE---LKEQWMAFLGAEKQ-----LQAEKALEHL-----EWALK 437
 QY 584 DOEKMEDEMKKRHHHEIFDLKEFDEALEQMTK 617
 DB 438 MTGAQMTLEQEHIMPVHKLREFFEQCAESRPE 471

RESULT 13

S00084
 myosin heavy chain, fast skeletal muscle - rabbit (fragment)
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 07-Sep-1990 #sequence_revision 07-Sep-1990 #text_change 20-Jun-2000
 C/Accession: S00084; S14807
 R/Maeda, K.; Sczakiel, G.; Wittlinghofer, A.
 Eur. J. Biochem. 167, 97-102, 1987
 A/Title: Characterization of cDNA coding for the complete light meromyosin portion of a
 A/Reference number: S00084; MID:87304245; PMID:3305014
 A/Accession: S00084
 A/Molecule type: mRNA
 A/Residues: 1-676 <MAE>
 A/Cross-references: EMBL:X05958; NID:g1622; PID:CAA29391.1; PID:g1364242
 A/Note: The sequence from Fig. 5 is inconsistent with that from Fig. 3 in having 561-Arg
 R/Maeda, K.; Roessch, A.; Maeda, Y.; Kalbitzer, H.R.; Wittlinghofer, A.
 FEBS Lett. 281, 23-26, 1991
 A/Title: Rabbit skeletal muscle myosin. Unfolded carboxyl-terminus and its role in molec
 A/Reference number: S14807; MID:91200294; PMID:2015900
 A/Accession: S14807
 A/Status: preliminary
 A/Molecule type: protein
 A/Residues: 668-676 <MA2>
 C/Superfamily: myosin heavy chain, myosin motor domain homology
 C/Keywords: actin binding; ATP; coiled coil; muscle; skeletal muscle
 F;21-676/Domain: light meromyosin <LME>

Query Match 5.2%; Score 173; DB 2; Length 676;
 Best Local Similarity 21.1%; Pred. No. 0.052;

Matches 124; Conservative 91; Mismatches 184; Indels 188; Gaps 30;

QY 141 ARGGSAGHTAVQEPFVEDVDVNASSEEND---SDALDGS---DDDLASDDYDSVSQK 193
 DB 43 SRGSG-----QAFYQOIEGLKRLQLEETKAKSALAHALQSSRRDDLLAEQYEEQEA 95
 QY 194 SHGRKQNKQKFKFPLSDLSLSTEQINEPQRMHCAPACQNGPAGIWMYMLPHLAHARTK 253
 DB 96 AEIQRAMSK-----ANSEVSQWR-TKCT-----DAIQTEELEKCKK 133
 QY 254 GARRVKLHRELAELEKDLQMRGASVTPCGEIVGQWKGLEDEKDYETVWPVMTIMNT- 312
 DB 134 LMQRLDADAEHVAVNS-----KCSLEKTKQRLQNEADL-----MIDVERSN 177
 QY 313 ---RLDKDNDKWLGMGNOELLETFDKYALRAHSYSGPQGRGMSVLMFE-----SSA 363
 DB 178 ATCARMDKKQRN-----FDKVLAEWKHKYBETOAELEASQKESRSLSTEVFYKKNAYEES 232
 QY 364 TQVLEA-----EHLHRELAMGMDRIAMG-----QKRSMSGGVRLQYGF 493
 DB 233 LDHLETLTKENKVLQOEISDL-TEQIAESAKHIELKVKKQIQOESESLQALAELEGES 291
 QY 404 LATKQDLDLFNQSQK-TRLEKELKSYQEMVVKELQISDNQOQLNRYFNKLSKONKHA 462
 DB 292 L-----EHEBKIRIQIQLNQLVQSEIDRKIAKDEBEIDQL-----KCN-HL 332
 QY 463 KYLEESLEIMSEKLRRTAEEDNRIVRTQKQHEQNEEMDAH--DRPFMDSIK----- 513
 DB 333 RVEVSGQSTLDAEIR---SRNDALRIKKKMGDLNEMEIQLNHNARQAAIAIKLNTQSG 389
 QY 514 ---QIH-----ERR---DAK-EENFEMLOOQERAKVYGOO--- 542
 DB 390 ILKDTQLHLDVAVGDDHKEQLAMVERKNMLQALIEELRASLEQTERSRVADDLLD 449
 QY 543 ---QINPSSNDCKRAAEVSSFFIEQEKMEEFVEE----- 577

DB 450 ASERVQLLHTON---TSLINTKKKLETDISOI---QGMEDIVQEARNAERAKKAITDA 503
 QY 578 ---REMLTKDOEKK--MEDMKKRHHHEIFDLKEPPEALEQMTK 619
 DB 504 AMMAEELKKEODTSAHLEKRMKQMEQTVKDLQORDEA-EQALAKGG 549

RESULT 14

B70356
 chromosome assembly protein homolog - Aquifex aeolicus
 C/Species: Aquifex aeolicus
 C/Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 02-Jun-2000
 C/Accession: B70356
 R/Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ove
 V.
 Nature 392, 353-358, 1998
 A/Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.
 A/Reference number: A70300; MID:98196666; PMID:9537320
 A/Accession: B70356
 A/Status: preliminary; nucleic acid sequence not shown; translation not shown
 A/Molecule type: DNA
 A/Residues: 1-1156 <AGP>
 A/Cross-references: GB:AE000699; NID:g2983238; PIDN:AA06839.1; PID:g2983243; GB:AE00065;
 A/Experimental source: strain VFS
 C/Genetics:
 A/Genes: xcpC
 C/Superfamily: chromosome segregation protein SMCI

Query Match 5.2%; Score 172.5; DB 2; Length 1156;
 Best Local Similarity 19.9%; Pred. No. 0.11;

Matches 82; Conservative 76; Mismatches 139; Indels 115; Gaps 15;

QY 268 LEKDLQMRGASVTPCGE-----YQGMKGL-----GDEKDYETVWPVMTI 308
 DB 555 LAKRMNLGRFSFPLNRRVVERPRLYPTTKGAVPAVNLVEYDPFKFKVYKVFQDTLI 614
 QY 309 INMTRDKDNDKWLGMGNOELL---EYFDKYEALRAHSYSGPQGRGMSVLMFESSAT 364
 DB 615 VENN-----ESAKAIGIGRYAVTLLEGELFEKSGV-----TGAV-----KSS 653
 QY 365 GYLAERLHRELAEWGLDRIAMGQKSMFSGGVRLQYGLATKQDLDLFNQSQKTRIK 424
 DB 654 GELNKKRYEEELQRLNAEBEKLNEBSIIQKIRIRHNLISEKTL-----LK 701
 QY 425 FELKSYQEMVVKELQISDNQOQLNRYFNKLSKONKHAUVLEESLEIMSEKLRRTAE-- 482
 DB 702 VSEKRIEELSGLQYEEK-----FKELNSKEYLILBEKLLNVEDKLELAEEIE 755
 QY 483 ---NRLVRQTKQHEQNEEMDAHDPFMDSIKQIHE-----RDKAKENPFM 528
 DB 756 YIEBKLNINLKLKGGDILKRRYSREGVAEKREYSKVRKQVSELEKSLNELERLANKTYEL 815
 QY 529 ---LQOERAKVYGOOQOINPSSNDCKRAAEVSSFFIEQEKMEEFVEREMLI 582
 DB 816 EYLEKEIQEKER-----EYLTERISLKKKEINLILFKKTKQE-VKEAEVAV 864
 QY 583 KQOEKMEDEMKK-----RHHEIFDLKE---FDALQOL 614
 DB 865 YVYIKQKELEKEIINLAKSKLGLKIKELKEKFEKEKKNLKVLEKIENTL 916

RESULT 15

S28589
 trichohyalin - rabbit
 C/Species: Oryctolagus cuniculus (domestic rabbit)
 C/Date: 12-Mar-1993 #sequence_revision 01-Mar-1996 #text_change 22-Jun-1999
 C/Accession: S28589
 R/Pietz, M.J.; Rogers, G.E.
 submitted to the EMBL Data Library, December 1992
 A/Description: Examination of the gene encoding rabbit trichohyalin.
 A/Reference number: S28589
 A/Accession: S28589

A: Molecule type: DNA
A: Residues: 1-1407 <F1B>
A: Cross-references: EMBL:219092; NID:91746; PID:CAA79519.1; PID:91747
C: Comment: Trichohyalin is a protein of the medulla of the hair and of the inner root sheath.
C: Covalent modifications to this protein include conversion of arginine to citrulline and
C: Genetics:
A: Introns: 46/3
C: Superfamily: trichohyalin; calmodulin repeat homology
C: Keywords: calcium binding; citrulline; EF hand; hair; tandem repeat
F: 49-81/Domain: calmodulin repeat homology <EF2>

Query Match 5.2%; Score 172.5; DB 1; Length 1407;
Best Local Similarity 21.08; Pred. No. 0.13; Indels 141; Gaps 17;
Matches 87; Conservative 71; Mismatches 116;

```
QY 278 SVIPGGEIYGOM-----KGIGEDKDYEIWM-----PPMVIIMNTRLDKDN 319
Db 7 SIIDITIFNOYASHDCDGVKKKDKILDRFGAVLORPHDPETVDMLELDR-DS 65
QY 320 DKWLGMGNQELLEYFDKYEALRHSYGPQGHGMSVLMPESSATGYLAEERLHRELAE 379
Db 66 DGLVG-----FDEP-----CLIFKLAQAAYVALGQ-----AS 93
QY 380 GLDRIAMGQKRSFSGVRLYGFLATKODLIFNHSQKTRL-----KTELKS 429
Db 94 GLDE-----EKRS-----HGEKGRLQNRQEDQRRELKD 125
QY 430 YQ-----EMVVKELRQISEDNQQLNFKNLASKO--NKAQVLEESLE----- 470
Db 126 RQFDEPERRRWQOEERELAEERERKRRERFQHSRYRDKERQRLQREERRAE 185
QY 471 -----IMSEKLRTAEDNRIVRRTKQHEQNREEMDAHDHFFMDSIKQI 515
Db 186 EQQLRRRKRDSEEFIEEQLR--REQQLKRELREERQORRRERBOHERALQEEEOQL 243
QY 516 HERRDAKEENFEMLOQOEERAKVGGQONINPSSNDCKRAEVSFIEFOE-----KEM 571
Db 244 LQQRKWRERPEEQOQLRELEIRERERQLEERERQOQLRERQLEERERERQOQLREL 303
QY 572 EEVFERERMLIK-----DOEKKMEDMKRHHHEIFDLEK--EFDLEALQIM 615
Db 304 EE-IRERERQLEERERERQLEERERERQOQLKRELIEIRERERQLEERERERQOQL 357
```

Search completed: April 6, 2004, 19:35:24
Job time : 24 secs